



## SUMMER INTERNSHIP

---

# Developping a Smart ChatBot - "GalaBot"

---

*Galactech Studio*

*Developed By:*

Kandil AMAL  
Smari MARIEM

*Framed By :*

Sir.Bouattour Mustapha

2020- 2021

# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Project Context</b>	<b>2</b>
1.1 Description of the project . . . . .	2
<b>2 Realisation</b>	<b>4</b>
2.1 Technologies . . . . .	4
2.1.1 Keras . . . . .	4
2.1.2 TensorFlow . . . . .	4
2.1.3 SpaCy . . . . .	4
2.1.4 Selenium . . . . .	5
2.1.5 Googletrans . . . . .	5
2.2 Abandoned methods . . . . .	5
2.2.1 LSTM . . . . .	5
2.2.2 Bert . . . . .	6
2.2.3 Dialogflow . . . . .	6
2.2.4 Ktrain library . . . . .	7
2.3 Methodology . . . . .	7
2.3.1 Project understanding . . . . .	7
2.3.2 Data Collecting . . . . .	7
2.3.3 Building the model . . . . .	7
2.3.4 Deploying . . . . .	10
<b>Conclusion</b>	<b>11</b>
<b>Bibliography</b>	<b>11</b>

# List of Figures

2.1	Keras Logo . . . . .	4
2.2	TensorFlow Logo . . . . .	4
2.3	SpaCy Logo . . . . .	4
2.4	Selenium Logo . . . . .	5
2.5	LSTM . . . . .	5
2.6	Bert . . . . .	6
2.7	Dialogflow . . . . .	6
2.8	Process flow diagram . . . . .	8
2.9	Screenshot English Question . . . . .	9
2.10	Screenshot Arabic Question . . . . .	9
2.11	Screenshot French Question . . . . .	9
2.12	Screenshot Arabic Question . . . . .	10

# Introduction

The first chatbot was developed in 1966 at MIT and they called it ELIZA. ELIZA, the mother of all chatbots, answered some very simple decision tree questions.

They are used to reproduce powerful interactions with users, to aid business processes, to gain information from large groups, as a personal assistant among others. Bots are also used by search engines to lag the web and archive new pages for future search.

A chatbot is digital with text and messaging or voice based applications. They help different groups of people or individuals to put their inquiries via text or voice.

In fact, our summer internship was in this context of building a Smart ChatBot for Galactech Studio that was suggested by its RD Department.

This internship allowed us to dive in the Natural Language Processing (NLP) cause ChatBots are concerned specially with that field. NLP is what allows chatbots to understand our messages and respond appropriately. So, without NLP, AI that requires language inputs is relatively useless.

There's no doubt that a lot of researches were taken to understand the whole process and a precise commitment to finalize it.

The purpose of this present report is to define and describe the work carried out throughout the project of implementing a Smart ChatBot.

This report will be divided in two chapters:

- In the first chapter, we are going to describe the project.
- In the second chapter, we will enumerate the technologies used in the whole project. In addition, we will define the different methods tested and the best alternative that we did take. To end with defining the steps that we worked with.

# Chapter 1

## Project Context

### 1.1 Description of the project

A chatbot or conversational agent is a software that can communicate with a human by using natural language. One of the essential tasks in artificial intelligence and natural language processing is the modeling of conversation. Since the beginning of artificial intelligence, it's been the hardest challenge to create a good chatbot. Although chatbots can perform many tasks, the primary function they have to play is to understand the utterances of humans and to respond to them appropriately.

The main task for our summer internship was building a smart ChatBot for different platforms to introduce the company, answer FAQs, offer customer services and technical help for our end users of the company. If the chatbot didn't recognize the question, it will send that question to the official page of Galactech's Facebook searching for an answer from a worker.

But, we have to mention the 5 moments to respect while building a chatbot:

1. Moment 1: Greeting
2. Moment 2: Listen
3. Moment 3: Acknowledgment
4. Moment 4: Answer
5. Moment 5: Unhappy path

1/

- Greet the user by their name
- Communicate the value your bot provides
- Give clear directions on how to interact with your bot
- Prompt the user on the next steps

2/

- Train your bot to recognize variations of the same user input
- Prepare for non-verbal and non-text inputs like emojis and images
- Teach your bot to understand profanity, positive and negative feedback, and chit chat

3/

- Identify key pieces of data captured from the user input that will affect the bot's answer
- Acknowledge what the bot heard before offering an answer so people know what the answer is based off of
- Allow users to change one or all pieces of data to affect the final results

4/

- Inject data and sources that influenced the bot's answer
- Explain why the bot thinks the results it offers are relevant to the person
- Offer reasons to believe the bot
- Give enough data for people to easily make a decision

5/

- Bot errors: Create specific responses for different types of issues that the business is responsible for. These can

---

be system errors, unsupported user asks, unsupported language, and limitations around the artificial intelligence of the bot. Each of these scenarios are moments to be transparent and offer guidance.

# Chapter 2

## Realisation

### 2.1 Technologies

#### 2.1.1 Keras

This is an open source software library that focuses on simplifying the creation of deep learning models. Written in Python, it can also be deployed on top of many other AI technologies such as Theano and TensorFlow. It runs optimally on both CPUs and GPUs, plus it is known for its user-friendliness as well as fast prototyping.



Figure 2.1: Keras Logo

#### 2.1.2 TensorFlow

With the initial release of this open source machine learning framework being 2015, it has been deployed across many different platforms and is easy to use. Created by Google at first, now all the top tech giants such as eBay, Dropbox, Intel and Uber use it extensively. With the help of flowgraphs, one can develop neural networks.



Figure 2.2: TensorFlow Logo

#### 2.1.3 SpaCy

It is an open-source library for Natural Language Processing (NLP) in Python language. The library is designed specifically for developers to build interactive NLP applications, which can process and ‘understand’ large volumes of text. spaCy can also be utilized to create information extraction or NLU systems, as well as pre-process text for deep learning.



Figure 2.3: SpaCy Logo

### 2.1.4 Selenium

Selenium Python bindings provides a simple API to write functional/acceptance tests using Selenium WebDriver. Through Selenium Python API you can access all functionalities of Selenium WebDriver in an intuitive way.



Figure 2.4: Selenium Logo

### 2.1.5 Googletrans

is a free and unlimited python library that implemented Google Translate API. This uses the Google Translate Ajax API to make calls to such methods as detect and translate.

## 2.2 Abandoned methods

### 2.2.1 LSTM

Long Short Term Memory networks (LSTMs) are a special kind of RNN, capable of learning long-term dependencies.

LSTMs are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods of time is practically their default behavior, not something they struggle to learn!

All recurrent neural networks have the form of a chain of repeating modules of neural network. In standard RNNs, this repeating module will have a very simple structure, such as a single Tanh layer.

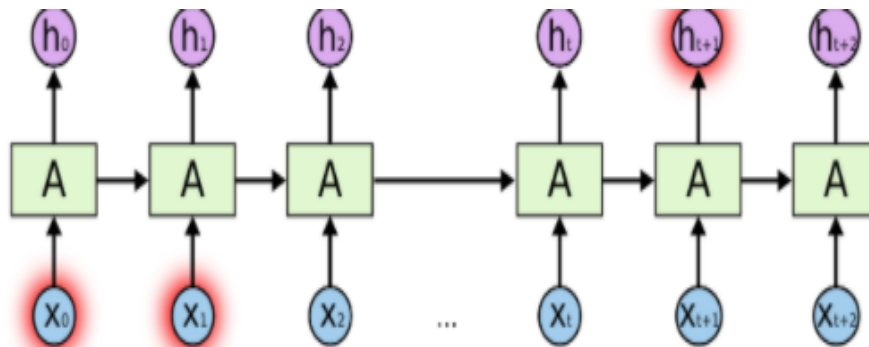


Figure 2.5: LSTM

This method takes a lot of time more than 12 for training. With the resources we have and the restriction in Google Colab we didn't manage to make it work.



## 2.2.2 Bert

Bidirectional Encoder Representations from Transformers (BERT) is a technique for natural language processing (NLP) pre-training developed by Google. Google is leveraging BERT to better understand user searches. BERT model comes with two pre-trained general types : the BERT-BASE model, a 12-layer, 768-hidden, 12-heads, 110M parameter neural network architecture, and the BERT-LARGE model, a 24-layer, 1024-hidden, 16-heads, 340M parameter neural network architecture; both of which were trained on the BooksCorpus[4] with 800M words, and a version of the English Wikipedia with 2,500M words.

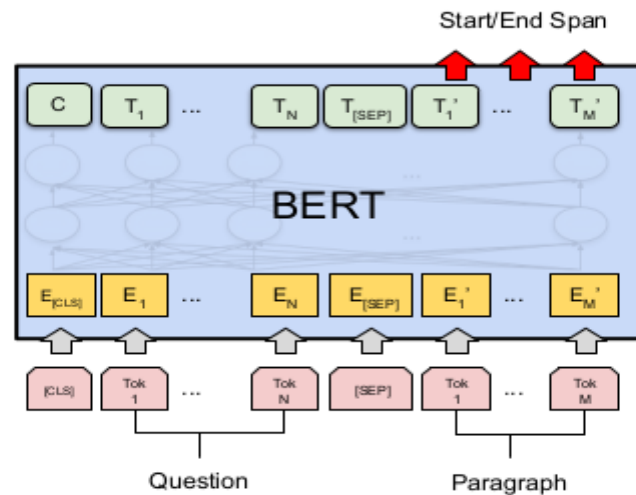


Figure 2.6: Bert

It is difficult to fine-tune Bert because our data is not structured the same way SQUAD was structured.

We can't work with the classic Bert because you have to provide a question and a context which is not the task we want to validate.

## 2.2.3 Dialogflow

Dialogflow is a natural language understanding platform used to design and integrate a conversational user interface into mobile apps, web applications, devices, bots, interactive voice response systems, and so on.



Figure 2.7: Dialogflow

This platform is based on intents that allows us to understand what the user wants and match it into an intent (action) of our chatbot. In our case we cant use intents because we have different questions and we cant regroup questions together.

---

## 2.2.4 Ktrain library

ktrain is a lightweight wrapper for the deep learning library TensorFlow, Keras and other libraries to help build, train, and deploy neural networks and other machine learning models. Inspired by Machine Learning framework extensions like "fastai" and "ludwig", ktrain is designed to make deep learning and Artificial Intelligence more accessible and easier to apply for both newcomers and experienced practitioners.

ktrain provides support for applying many pre-trained deep learning architectures in the domain of Natural Language Processing and BERT is one of them. To solve this problem, we will be using the implementation of pre-trained BERT provided by ktrain and fine-tune it to answer the questions.

The model used with Ktrain provided a score with each answer to a certain question and we have to pick the answer with a higher score. But, the main problem with Ktrain was the time spent in the execution to make a result.

## 2.3 Methodology

### 2.3.1 Project understanding

First started first by analyzing the project: What ? why ? How ?

We understood what we are asked to do: Building a smart chatbot/ we checked different types of chatbots and which one correspond to our mission better like hybrid chatbots. we choose the technologies that we will use and type of models we will try.

### 2.3.2 Data Collecting

We collected data from different gaming websites and social platforms such as Galactech's Facebook page and the company provided us with a small excel file containing some questions and their answers.

### 2.3.3 Building the model

We tried many models such as LSTM, Bert, Ktrain. And smart platforms like Dialogflow to choose at the end the easiest solution which is using ChatterBot

#### ChatterBot

ChatterBot is a Python library that makes it easy to generate automated responses to a user's input. ChatterBot uses a selection of machine learning algorithms to produce different types of responses. This makes it easy for developers to create chat bots and automate conversations with users.

#### How ChatteBot works ?

ChatterBot is a Python library designed to make it easy to create software that can engage in conversation.

An untrained instance of ChatterBot starts off with no knowledge of how to communicate. Each time a user enters a statement, the library saves the text that they entered and the text that the statement was in response to. As ChatterBot receives more input the number of responses that it can reply and the accuracy of each response in relation to the input statement increase.

The program selects the closest matching response by searching for the closest matching known statement that matches the input, it then chooses a response from the selection of known responses to that statement.

And here in the image below, the process of ChatterBot is illustrated with details:

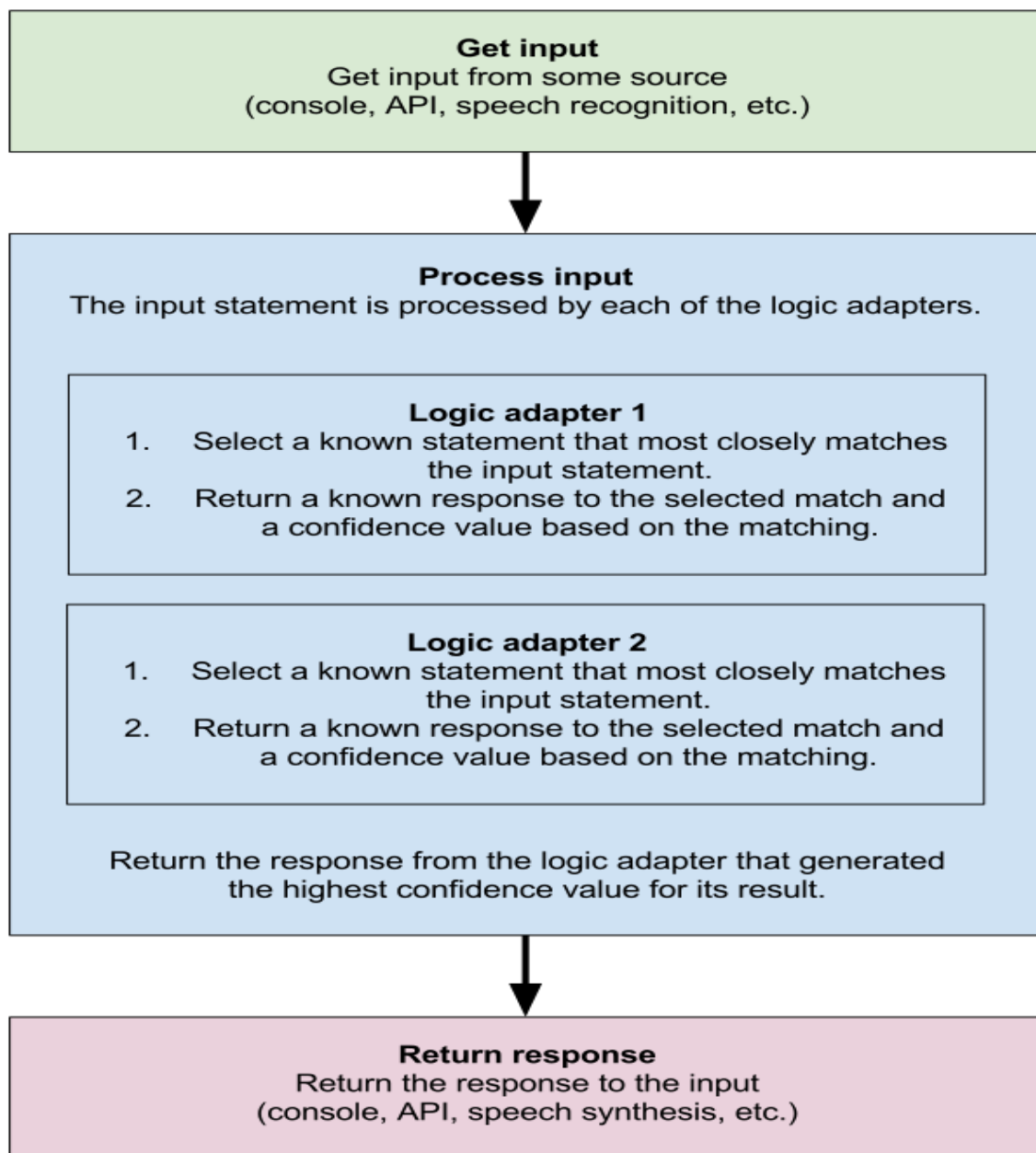


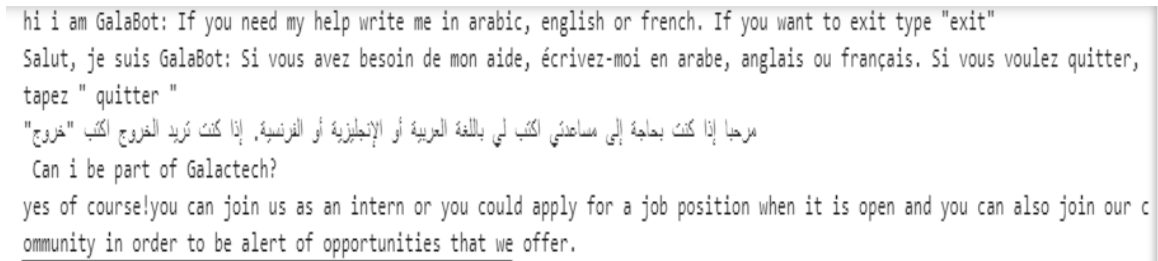
Figure 2.8: Process flow diagram

## Installation

It is recommended that we start by installing the latest version from the Python Package Index (PyPi).

## Screenshots for the results

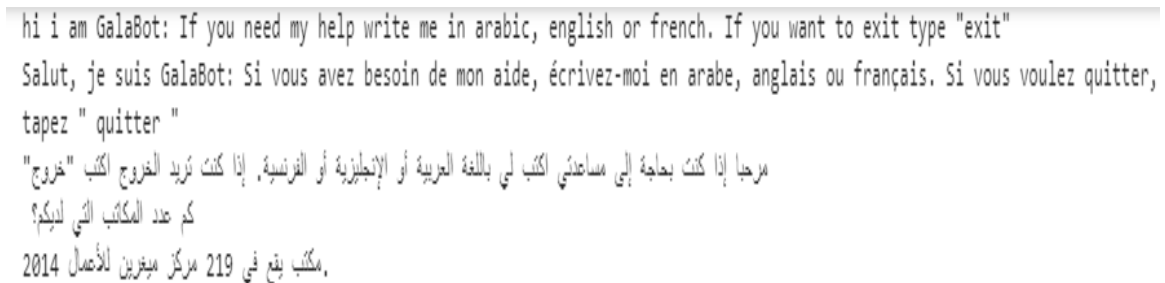
The photo below, shows how the chatbot did answer for a question asked in English.



hi i am GalaBot: If you need my help write me in arabic, english or french. If you want to exit type "exit"  
Salut, je suis GalaBot: Si vous avez besoin de mon aide, écrivez-moi en arabe, anglais ou français. Si vous voulez quitter, tapez " quitter "  
مرحبا إذا كنت بحاجة إلى مساعدتي اكتب لي باللغة العربية أو الإنجليزية أو الفرنسية. إذا كنت تريد الخروج اكتب "خروج"  
Can i be part of Galactech?  
yes of course!you can join us as an intern or you could apply for a job position when it is open and you can also join our c  
ommunity in order to be alert of opportunities that we offer.

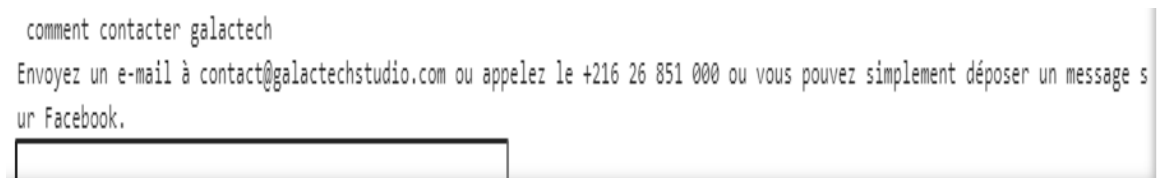
Figure 2.9: Screenshot English Question

Therefore, the chatbot can translate the question in three different languages such as English, French or Arabic. So, it could fit in order to answer the questioner in the same language that he asked. As explication, the next two images shows that.



hi i am GalaBot: If you need my help write me in arabic, english or french. If you want to exit type "exit"  
Salut, je suis GalaBot: Si vous avez besoin de mon aide, écrivez-moi en arabe, anglais ou français. Si vous voulez quitter, tapez " quitter "  
مرحبا إذا كنت بحاجة إلى مساعدتي اكتب لي باللغة العربية أو الإنجليزية أو الفرنسية. إذا كنت تريد الخروج اكتب "خروج"  
كم عدد المكاتب التي لديكم؟  
مكتب يقع في 219 مركز ميغرين للأعمال 2014

Figure 2.10: Screenshot Arabic Question



comment contacter galactech  
Envoyez un e-mail à [contact@galactechstudio.com](mailto:contact@galactechstudio.com) ou appelez le +216 26 851 000 ou vous pouvez simplement déposer un message s  
ur Facebook.

Figure 2.11: Screenshot French Question

---

But, if the chatbot can't answer the question, it will be sent to Galactech's Facebook page as a message so that an agent can communicate with that client.



Figure 2.12: Screenshot Arabic Question

#### 2.3.4 Deploying

Using Node.js, this chatbot was supposed to be implemented in different platforms such as Galactech's website, Facebook page, Borgdena platform but because of the pandemic this task was not realized.

# Conclusion

In this report, we have represented in first place a description of the project then we moved to realization where we talked about technologies , abandoned methods and methodology.

Thanks to our experience in Galactech Studio, we learned from a personal point of view how to manage our project in a methodical and organized way.

This work has been beneficial to us as it allowed us to put into practice our theoretical knowledge acquired throughout our training at ESPRIT School of engineering.

It also allowed us to deepen our knowledge and learn new technologies. Our project is therefore a source of technical, cultural, personal and human enrichment.

Unfortunately, due to the pandemic, we couldn't finish the final part of this project which is deployment and we couldn't learn how to use Node.js in our task but luckily most of the work was done.

# Bibliography

- [1] *The official SpaCy's Site*,  
<https://spacy.io>
- [2] *The official Googletrans's Site*,  
<https://pypi.org/project/googletrans/>
- [3] *Defining the technologies*,  
<https://spacy.io>
- [4] *The official Site for Selenium*,  
<https://selenium-python.readthedocs.io/installation.html>
- [5] *The libraries for Chatbot*,  
<https://analyticsindiamag.com/top-python-libraries-for-chatbot-development/>
- [6] *Description for Dialogflow*,  
<https://chatbotslife.com/getting-started-with-dialogflow-and-building-my-first-bot-newbies-guide-d025d4eed3b2>
- [7] *Dialogflow*,  
<https://en.wikipedia.org/wiki/Dialogflow>
- [8] *Defining Bert* ,  
[https://en.wikipedia.org/wiki/BERT\\_\(language\\_model\)](https://en.wikipedia.org/wiki/BERT_(language_model))  
<https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1194/reports/default/15848021.pdf>
- [9] *Presenting the steps for building a ChatBot* ,  
<https://medium.com/@jseijs/how-to-build-your-own-nlp-for-chatbots-9b4c08eb03a9>
- [10] *Visited Githubs*,  
<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>  
<https://github.com/Joeyipp/chatbot-seq2seq>