A blue and black logo

Description automatically generated  
  
  
  
  
  
  
Chat Bot

Table of Contents

[Objectives: 2](#_Toc156302456)

[How to train models based in personal information 3](#_Toc156302457)

[What is Langchain going to be used for? 3](#_Toc156302458)

[Models available to train and deploy application: 3](#_Toc156302459)

[Using Gpt4all 3](#_Toc156302460)

[Using Bard API 4](#_Toc156302461)

[Dialog flow: 4](#_Toc156302462)

[Free open-source alternatives: 6](#_Toc156302463)

[Vertex AI: 6](#_Toc156302464)

[What is Vertex AI: 6](#_Toc156302465)

[Using ChatGPT API 8](#_Toc156302466)

[Gpt3 8](#_Toc156302467)

[GPT4 8](#_Toc156302468)

[Structure Diagram: 9](#_Toc156302469)

# Objectives:

This document purpose is to analyze the viability and the options for Moro technology develop a chatbot able to be integrated with google chat.  
The Chat bot should have the following features:

* Be able to be trained and retain the knowledge of pdf’s containing company information.
* Be developed in a way that is possible to be integrated with third party applications (ex: TravelPerk)
* Possibly be able to recognize information about the user and give confidential information about him (ex: display salary or available trip budget)

Further features need to be discussed with the HR department.

# How to train models based in personal information

To be able to train the chatbots form pdf’s documents a tool called Langchain comes handy.  
  
What is Langchain:  
Langchain is an open-source Python library designed to facilitate the development of applications that integrate language models like GPT-3 or GPT-4 into their core functionality. It provides tools and abstractions that simplify the process of building complex applications that leverage the capabilities of large language models.

### What is Langchain going to be used for?

Langchain is going to be used for splitting the pdf’s in to chunks, converting does chunks in to vectors and to do Similarity Search.

Models available to train and deploy application:

## Using Gpt4all

GPT4All is an open-source software ecosystem that enables users to train and deploy powerful and customized large language models (LLMs) on everyday hardware. It is designed to run locally on consumer-grade CPUs, making it completely cost free. This also ensures that the data of Moro technology will not be used or stored in the cloud by Open AI or Google.

**Note the following models are not local called by using a API key meaning the data will not be confidential:**

## Using Bard API

Gemini Pro

new release this year (2024):

* Gemini Ultra:
* Gemini Nano (for Android AiCore):

Note: Gemini Api Is **not** available in EU and UK, making the only alternative possible to use Vertex AI

A screenshot of a computer

Description automatically generated

## Vertex AI:

What is Vertex AI:  
  
The managed machine learning platform Vertex AI, created by Google Cloud, simplifies the process of creating, honing, and implementing AI models. It makes it simpler for data scientists and engineers to manage AI projects from beginning to end in a single environment by integrating multiple Google Cloud services for AI into a single user interface and API. To improve model building and training efficiency, Vertex AI provides a number of tools and features, such as Tabular Workflows and the AI Training Reduction Server. The platform's goal is to make AI more accessible to a wider variety of users by streamlining the deployment of AI models and supporting sophisticated machine learning operations.  
  
Pricing of Vertex AI:  
 $ 1,000 in FREE trial credits are awarded for use with Vertex AI Search and Conversation. These credits are valid for 1 year period beginning from account signup with Vertex AI Search and Conversation starting on Jul 18, 2023.  
A screenshot of a computer

Description automatically generated  
A screenshot of a computer

Description automatically generated

## Using ChatGPT API

None of the chat GPT api is free, having the following prices:

“Multiple models, each with different capabilities and price points. Prices are per 1,000 tokens. You can think of tokens as pieces of words, where 1,000 tokens is about 750 words. This paragraph is 35 tokens.”

### GPT3

A screenshot of a computer

Description automatically generated

### GPT4

A screenshot of a computer

Description automatically generated

# Structure Diagram:

A diagram of a company

Description automatically generated

Dialog flow:  
Dialogflow, formerly known as Api.ai, is a cloud-based software developed by Google that provides natural language understanding (NLU) and dialogue management functionalities. It's designed to help developers build conversational interfaces for websites, mobile applications, popular messaging platforms, and IoT devices.  
Some key features include:

* Natural Language Understanding (NLU): Dialogflow enables developers to construct conversational interfaces that can comprehend context, manage conversations, and respond appropriately by interpreting user input (text or voice) and understanding its meaning.
* Integrations: It may be integrated with a number of systems, including Slack, Google Assistant, Amazon Alexa, Facebook Messenger, and more. This enables the multichannel deployment of chatbots and other conversational interfaces by developers.

There are 2 different types of agents in **Dialogflow**:  
A screenshot of a computer

Description automatically generated

Free open-source alternatives:  
  
Rasa:   
An open-source machine learning framework for building AI-powered chatbots and voice apps. Rasa is known for its robust, context-based conversational experiences, utilizing Natural Language Understanding (NLU) and Machine Learning (ML)

Botpress:   
An open-source conversational AI software, which allows you to build chatbots using visual flows and small amounts of training data

Wit.ai:  
An open-source chatbot builder owned by Meta, suitable for creating conversational interfaces for various applications, including Facebook Messenger. It uses natural language processing (NLP) and machine learning algorithms​  
  
Botman:  
A PHP library for chatbot creation, which supports multiple messaging platforms like Slack, Telegram, and Facebook Messenger. It requires programming skills and is framework agnostic​