# CHAPTER ONE

# INTRODUCTION

## 1.1 Background to the Study

Coronavirus disease 2019 is an illness caused by a new coronavirus called Sever Acute Respiratory syndrome coronavirus-2(SARS-CoV-2) (WHO, 2019). It was first found in Wuhan City, Hubei Province China. On March 11, 2020, World Health Organisation declared it a global pandemic as the number of infected countries grows (Espinoza, Crown, and Kulkarni 2020) A few months later, it was detected in more than 70 locations internationally, including Nigeria (WHO, 2019). This has imposed a global threat due to its highly contagious nature, which has affected all sectors of the world economy and society. All the healthcare workers caring for patients with the disease were at elevated risk of exposure and also, close contacts of persons with covid-19.

Dada, 2021 reported that since the outbreak of COVID-19, numerous preventive and control measures have been applied globally to contain the disease.

Omaka-Amar et.al, also pointed that report on the disease from December 2019 to 30th April 2020 shows a continuous worldwide increase with a total of 3,090,455 confirmed cases, 217,769 fatalities, and 1,007,971 discharged cases distributed across 210 countries.

In Nigeria, the first outbreak of COVID-19 occurred on 27th February 2020 through an Italian businessman who visited the country Nigeria, from then, government began the battle at that moment with the help of other health agencies targeted at curtailing the spread of the disease. Amazingly, despite all preventive and control efforts following the outbreak, the disease as of 30th April 2020 had spread to 36states with 1,932 confirmed cases, 319 discharged cases, and 58 deaths (Omaka-Amari et al., 2020)

The challenge was how to stop the widespread of the disease without affecting the activities of people around the world. In the hospitals, the crowded waiting rooms could make the problem worse, therefore, limiting the population is going to be key to managing the community spread of the disease and engaging them on their own to find out their health status up to 50% before seeking the attention of a physician. This challenge has also put many countries into action in search of solutions.

Therefore, with the emergence of digital technologies, Chatbots evolved at an unprecedentedly rapid pace in almost all key sectors, including healthcare, hospitality, and many other important sectors(Natsheh & Jabed, 2022). No doubt that this will go a long way in mitigating the problem. One of the researchers who have developed a similar Chatbot in the past (Dada, 2021), named “I SABI COVID-19 CHATBOT”, used Dialogflow, a cloud-based framework by google and their aim was only to spread awareness of the coronavirus. The limitation to the final output was inability to make referrals to emergency services provided by the hospitals and clinics. Lei et al. (2021) also proposed a covid-19 chatbot known as “COVID-19 Smart Chatbot” for Ameriac using Rasa frame work but the functionalities only capture their terain. However, this study will improve on their Chatbot by focusing on functionalities and scope improvement by using an open-source framework with code flexibility. To actualize this, A Covid-19 AI-Based Chatbot will be proposed with the use of Python 3 and Rasa framework. The focal point will be the symptoms and the contacts made by the suspected patient. The symptomatology of the patient includes fever, malaise and, dry cough. It will be a public service, which will be free and available for both medical professionals and lay people.

## 1.2 Statement of the Problem

The fatal COVID-19 disease, which is brought on by the Corona Virus, is currently the subject of a major global war(Vanathi & SriPradha, 2020). It has dawn on human race that life is larger than logic, without life nobody will think of progress or war. Since Covid-19 was declared a pandemic (WHO,2020), many countries have been in search of solutions on how to prevent, diagnose, and treat the disease, but in all these, the death rate is still on the increase (Prianto & Harani, 2021). Nigeria as a country is inclusive; despite the effort of the Nigeria Center for Disease Control and Prevention, new cases and death rates have ravaged the country for long time. People find it difficult to adhere to the rule of regular washing of hands, facemask usage, social distancing, and movement restriction(Oladipo et al., 2020). The disease has claimed many lives due to a lack of health facilities, lack of health workers, and computer-based systems that can aid the masses(Omaka-amari et al., 2020). Lei et al , ( 2021) proposed covid-19 chat bot system for American with the use of Rasa framework but only captured American environment on its tracing and tracking location. Dada, (2021) proposed a chatbot system that will combat covid-19 in Nigeria but could not make a referral to the professional doctors and nurses provided by the hospitals and clinics for the purpose of an emergency situations. The above challenges have led to the proposal of a COVID-19 AI CHATBOT FOR NIGERIA that will be able to take the symptoms of the patients in Nigeria, location, contact, and other necessary details and come out with a robust functionality that can trigger an emergency responds to covid patient by NCDC and professional doctors. The benefit is that NCDC will be able to track and trace any individual with the covid-19 disease and give the patients immediate necessary medical attention hence, reducing the rate of infection.

## 1.3 Research Question

The following are pertinent research questions

1. How does the use of a COVID-19 chatbot increase the individual awareness and preventive measures in Nigeria?
2. Does the use of a COVID-19 chatbot contribute to a reduction in stress level among the medical practitioners and health works in Nigeria?
3. Does the use of COVID-19 chatbot cushion the effect of the shortage of specialists in Nigeria hospitals?

**1.4 Aim and Objectives**

The principal aim of the study is to develop Covid-19 AI Based Chatbot that will detect patients with covid-19 in Nigeria and channel them to the appropriate medical professional for quick intervention.

The specific objectives are to:

1. Develop AI based covid-19 chatbot with Rasa framework that will handle conversation between the user and chatbot agent.
2. Integrate Rasa framework with WhatsApp as the user interface.
3. Build a database of cases reported.

**1.5 Significance of the Study**

The study on the development of COVID-19 AI- BASED CHATBOT will be beneficial to all the health parastatals such as NCDC, WHO and Nigeria government at large. The information provided by the AI-Based Chatbot will help healthcare workers to locate and identify covid-19 suspected patients which will also make NCDC prepare and mobilize appropriate preventive and control activities in real time.

The study will go a long way to assist individuals who have the symptoms of covid-19 by giving them useful information on how to contact the appropriate authority before it gets too late. However, the study will help to cushion the effect of lack of medical personnel in different general hospitals in Nigeria by lifting their work load through attending to patients on simple issues.

**1.6 Scope of the Study**

Due to the crowd and temperature sensitivity of the Covid-19, the study is limited to Nigerian only. The conversation is only in English text and the chatbot is not an alternative to medical attention but rather a directive measure and awareness from the satisfied authorities of NCDC on how an individual can obtain a useful information about covid-19 and places to get medical attention

## 1.7 Operational Terms

## This section, will explain all the necessary terms and acronyms.

## Entity: This is considered useful information from the user input that can be extracted. The entity is the noun in the sentence that receives the action. For instance, if the user says “I want to consult a doctor”. The entity here is Doctor

## Intent: It is the target of the user input. For instance, if the user types “I want to consult a doctor,” The intent here is CONSULT.

## Action: This is an operation that can be performed by the bot. it could be replying to something in return, querying a database, or getting information from API.

## Stories: This is a sample interaction between the user and the bot.

## NLG: This stands for Natural Language Generation. It is part of the Rasa framework that responds to the user in natural language. Its main goal is to limit gaps in communication involving humans and machines.

## NLU: This stands for Natural Language Understanding. It is part of the framework that focuses on understanding the users’ messages or input.

## NLP: This stands for Natural Language Processing. Its goal is to convert the users’ text or speech data into structured data.

## DIET: this is known as Dual Intent and Entity Transformer. It is an inbuilt model in rasa that classifies user input intent and extracts the entity as the same time.

**CHAPTER TWO**

**LITERATURE REVIEW**

**2.1 Conceptual Framework**

Among all the architectural framework of NLP, this study will adopt the concept of Rasa Python Framework from(Patole et al., 2021) which is based on NLP (Natural Language Processing), NLU (Natural Language Understanding), and NLG (Natural Language Generation) models and they are branches of AI and ML working together to achieve a goal.

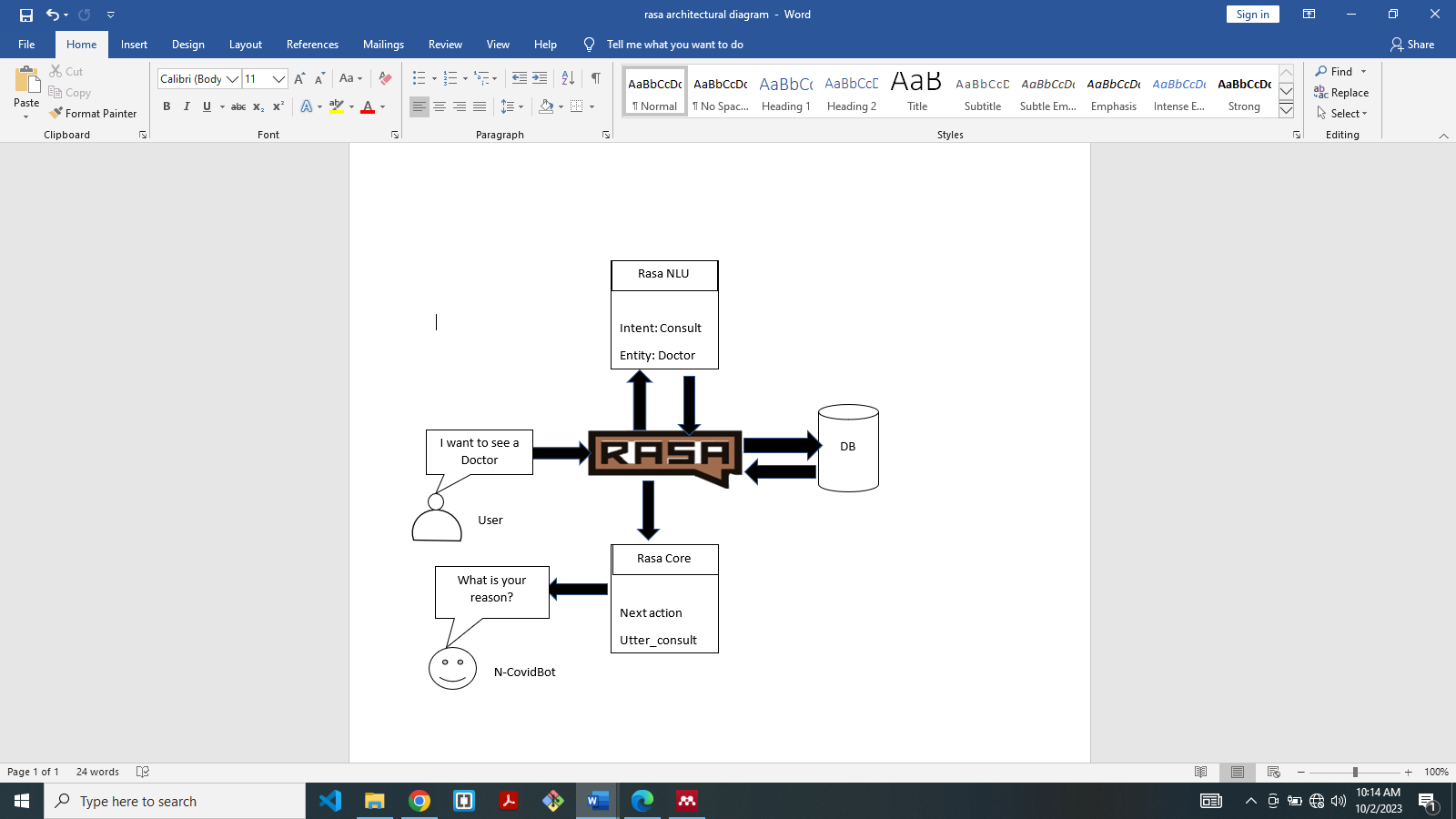
The figure 2.0 shows how NLU and NLG work together in a conversational AI Chatbot using Rasa

Figure 2.1: Conceptual Rasa framework (Patole et al., 2021)

From figure 2.0, the user will send input to the Rasa framework, then the Rasa NLU will separate the input into to intent and entity. For instance, the user input is “I want see a doctor”, then NLU used the inbuilt DIET to classify the intent as “consult” and extract the entity as “Doctor”. The intent is the verb while entity is the noun in the user input. Its main duty is to interpret the user input in a structured data format. Rasa core takes the structured input from the NLU and predict the next best action. For instance, the next action is to send responds to the user by the N-Covid Bot which is utter\_consult. Database backend is used to store reported cases from the user. NCDC staff has access control the database for tracking and tracing patient based on the report submitted.

## 2.2 **Empirical Review**

Developing AI chatbot is one of the ways people could be correctly informed and guided. Below are some authors who have ventured into solutions but there were some limitations with respect to population and methodology.

AI Covid-19 Awareness Chatbot was proposed by (Patil et al., 2021) for Indians for diagnostic evaluation and recommending immediate measures when a patient contract covid-19. Due to their advanced technology and robust database, they used a cloud-based framework known as Dialogflow, but this study discovered that the chatbot is customized only for Indians and the framework suitable for their database.

Dada (2021) proposed a chatbot for Nigeria named “I SABI” using DialogFlow as the framework from the Google cloud platform. Owning to the limited information about Nigeria in issues of covid-19, the chatbot could not meet up with some functionalities such as referrals to emergency services provided by hospitals and clinics. Therefore, to conquer the above limitation, a platform that will allow the developers to have full control of the code should be used such as the Rasa platform.

COVID-19 Smart Chatbot Prototype for Patient Monitoring was proposed by Lei et al. ,(2021) for the United State of America(USA). They used natural language processing and natural language understanding (model) that is embedded in the Rasa framework but the challenge is that the chatbot can only recognize USA terrains, restricting the function only for Americans.

However, from the above-related works, we will draw our main concern on the proposed Covid-19 AI-Based Chatbot for Nigerians from the following 1. Author Name, Chatbot Name, Country of the Chatbot, and Framework used

|  |  |  |  |
| --- | --- | --- | --- |
| Author Name | Chatbot Name | Country Origin  (Population gap) | Framework  (Methodology gap) |
| Vistro et al. 2021 | MedBOT | India | N-gram, TFIDF Technology, Dialogflow |
| Mittal et al. 2021 | Web-chatbot | Italy | GD, NLP, AND FFNN |
| Patil,Shelke,and Raut 2021 | AICovid-19 Awareness Chatbot | India | NLP, Dialogflow |
| Dada 2021 | I Sabi | Nigeria | NLP, Dialogflow |
| Lei et al. 2021 | COVID-19 Smart Chatbot | United State of America | NLP, NLU, RASA |

Table 2:1 Summary of Author Review on CHATBOT

**2.2.1 The Epistemology of COVID-19**

The short for Coronavirus Disease 2019 is COVID-19. It is an infectious disease that is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and was initially recognized in 2019 in Wuhan, China. The most common symptoms of COVID-19 include cough, fever, and shortness of breath. Other symptoms may include diarrhea, muscle pain, sore throat, sputum production, abdominal pain, and loss of taste and smell (Salman et al. 2020)

According to Salman et al (2020), COVID-19 typically spreads during close contact and via respiratory droplets produced when infected patients sneeze or cough. It may also spread through fomite transmission. For example, touching a fomite (contaminated surface) and then touching the body's mucous membranes, such as the mouth, nose, or eyes, could potentially introduce the pathogen into the body (Banjar et al., 2020).

The reality awareness of COVID-19 is undeniable, though, some Nigerians are still in doubt even as the International Health Regulations Emergency Committee of the WHO has declared that the outbreak is a “public health emergency of international concern”. This was declared on January 30th, 2020, and within 6 weeks, on March 11th WHO confirmed it a global pandemic (Espinoza et al., 2020). The declaration shows that the world has been trapped by the challenge of a deadly disease. This trigged a loud cry for help from every aspect of life to contain and mitigate the disease.

Owning to this, the Center for Disease Control and Prevention (CDC) then came up with interim guidance for healthcare facilities to address community transmission of COVID-19. The role was to reduce morbidity, and mortality and minimize disease transmission (Espinoza et al., 2020). This guidance cut across the whole world including Nigeria.

The first case of COVID-19 in Nigeria was identified on 27th February 2020 by a visiting Italian (Jo & Oj, 2020). Immediately, the Presidential Task Force on COVID-19 was constituted by President Muhammadu Buhari on March 9, 2020, to coordinate the country’s multi-sectoral, inter-governmental efforts to contain the spread and mitigate the impact of COVID-19(Adesegun et al., 2020).

Unfortunately, measures established by the task force to combat COVID-19 were faced with lots of challenges reneging from the political, social, economic, financial, and behavioral attitudes of Nigerians.

Research shows that the major challenge that undermined the control and prevention of COVID-19 in Nigeria was the issue of poor compliance attitude (Omaka-amari et al.,2020), and the lack of medical laboratory equipment for testing covid-19 in the 6 geopolitical zones. According to Uchejeso et al. (2020), the laboratory sites according to geopolitical zones are South West – 3 Laboratories (Lagos and Osun), South-South – 1 Laboratory (Edo), South East – 0 Laboratory, North Central/ FCT – 1 Laboratory, North West – 0 Laboratory and North East – 0 Laboratory. It could be deduced from the foregoing review that Nigeria is handicapped medically, financially, and professionally but with aid of AI-based systems, there will be hope.

However, some chatbot systems developed to combat with this deadly disease could not meet up with expected functionalities. Dada, (2021) proposed a chatbot for Nigeria to combat with covid-19 but ended in limited functionality. Lei et al, (2021) proposed a chatbot system for the same purpose but the scope was only within American.

**2.2.2. Mitigation with the Aid of AI chatbot**

The use of AI system in medical field has been in existence. This idea of thinking machines was proposed by an English Mathematician named Alan Turing. AI caught the attention of various scientists, computer enthusiasts, and even fiction writers after his proposal. Slowly, for many decades, society started to pursue more research on the feasibility of building machines that think like humans (Sathishkumar, 2020).

Today artificial intelligence has gained ground virtually in every field of life. One area of life where AI is making a significate impact is the pharmaceutical and

healthcare industries. It can address some critical issues such as the shortage of staff in the healthcare domain, the crisis in the skilled workers in the hospital, and easy access to new treatments, especially in the time pandemic like covid-19 (Espinoza et al., 2020). AI is the hub of machine learning and deep learning. Chatbot algorithm depends on Deep learn and machine learning algorithm.

**2.2.3 Chatbot**

According to Oxford English Dictionary, Chatbots are computer programs designed to simulate conversation with human users, especially over the internet. They are AI based conversational applications that enable users to communicate and interact with software (Almalki& Azeez ,2020). Also, Kate Bush saw it as a programming that simulates the conversation or "chatter" of a human being through text or voice interactions. In the advanced approach, a chatbot is an “online human-computer dialog system with natural language.(Darapaneni et al., n.d.) Therefore, a chatbot is simply a dedicated system created to interact with human beings in a natural form.

In 1966, Joseph Weizenbaum at MIT created the first chatbot that, arguably, came close to imitating a human called ELIZA. Given an input sentence, ELIZA would identify keywords and pattern-match those keywords against a set of pre-programmed rules to generate appropriate responses (Hutapean.d.). Ever since that moment, chatbot creation has gained ground in different areas of life endeavors, especially in health care and pharmaceutical area.

In line with this, Battineni, Chintalapudi, and Amenta they concurred that after the rise of the web and mobile apps, virtual chatbot applications are the latest invention of digital design.

The above-mentioned chatbots are dedicated to one country or the other. Most Covid-19 chatbots are built for a particular country due to some reasons such as terrain recognition, and temperature difference(Daniel & Bamidele, 2020). Some researchers suggested that cold and dry conditions can escalate the transmission while warm and humid conditions can slow it(Daniel & Bamidele, 2020). This denotes that countries with high latitudes will likely have higher cases and death. Therefore, is necessary for Nigeria to have her own dedicated COVID-19 CHATBOT for high precision.

2.2.4 **Trends of COVID-19 Occurrence in Nigeria**

Many Nigerians saw covid-19 pandemic as a Chinese and oversea disease. Some even attributed it to the punishment for their sins and consumption of what they called unclean animals. Despite the belief of many, the first case of COVID-19 in Nigeria occurred on 27 February 2020 in Lagos state (Omaka-amari et al., 2020). Many doubted it, but as time goes on, evidence began to surface.

Ozili( 2020), observed that surges of daily increase in the number of new cases began to occur from the 19 of March 2020 with the number of confirmed cases rising steadily to the emergency level. As of April 30, 2020, the total number of confirmed cases rose to 1,932 with 58 deaths and 319 discharged cases (Omaka-amari et al., 2020). Nigerians began to realize the reality on the ground. Some started believing that covid-19 can kill and some embraced it with mixed feelings calling it fake information.

According to Ohia (2020), the symptoms exhibited by COVID-19 disease range from fever, cough, respiratory symptoms, shortness of breath, and breathing difficulties. The infection affects all ages and equally affects both the healthy and people with low immune (Jo & Oj, 2020). They pointed that most affected adult patients are in the age bracket of 25 and 89 years with few cases occurring among children and infants in Nigeria. According to Jo and Oj (2020), the greater part of confirmed cases in Nigeria is in the age group 31-40 years (23%) with a predominance of males (69%) compared to females (31%). Statistics show that the majority of the infected patients have travel histories to high-risk countries while some were in contact with those infected patients with a travel history equally those without epidemiological links with those infected patients with travel history were also infected too.

As of the time of writing this project, National Obstetric Fistula Center, Abakalike in Ebonyi State (NOFIC) reported a new variant of covid-19 on 22nd August, 2023 that is more deadly and dangerous. They advised that crowded place, keeping 1.5distance, double face masks and washing of hands should be observed with diligent. (https://nofic.org.ng/). Therefore, there is an urgent need for more ideas that will help battle with the disease in Nigeria both now and in the future.

**2.2.5 Preventions and Drawbacks**

All the effort to prevent covid-19 have a great side-effect on the economy and the normal human life style because of the mode of infection which is more or less ant- social such as social-distance, wearing face-mask, and lockdown. People finds it challenging to observe the rules and this has led to a serious battle on how to mitigate the disease in the whole world.

In many developed countries, prevention of covid-19 is more scientific and systematic.

According to Daniela and Bamidele(2020), primary, secondary, and tertiary preventive measures as suggested are equally used sophisticated technology to mitigate and contain the spread. Buheji and Buheji (2020), also cited Chinese apps that was developed to monitor the movement of their citizens through colour codes instead of using police and the army to stop them. This approach helped them to save time, and resources and equally reduce the transmission rate.

Nigerians relied only on primary, secondary, and tertiary preventive measures to contain the covid-19. According to Daniel and Bamidele (2020), the primary prevention is more of public health measures which include personal hand hygiene, respiratory etiquette, environmental measures, social distancing, and travel-related measures. They describe secondary prevention as the strategies employed for early detection of the disease and prompt action using the available test facilities to identify the affected person. Then tertiary, is the point at which the patient is rehabilitated.

However, those measures need to be supported by AI technology to meet up with the global standard in Nigeria.

**2.3 Theoretical Framework**

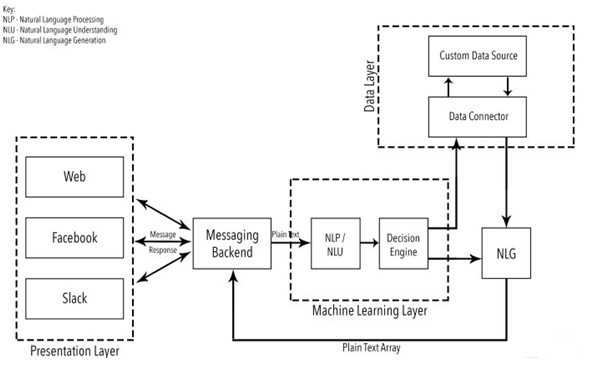
 The bedrock of all the AI conversational agent is rooted in NLP, NLU and NLG which are branches of AI. The fundamental technologies for chatbot are ML, DL, NLP and AI. Every AI based Chatbot frameworks such as Google DialogFlow, Python Rasa, Facebook Wit.AI, IBM Watson, Amazon Alexa Skill Kit and Microsoft Bot Framework depend on them. This study, will adopt the theoretical framework of AI based chatbot from (Ayanouz & Benhmed, 2020) displayed below

Figure 2.2: Theoretical Framework of AI Based Chatbot(Ayanouz & Benhmed, 2020)

The foundation of AI- based chatbot is NLP algorithm which has made it possible for input text to be understood, processed, and determine the next suitable action. In figure 2.2, there are three layers that determine the flow of plain text (input data).

Presentation Layer: It shows different UI channels (Web, Facebook and Slack) for the user to interact with the bot. The direction of arrow shows the movement of data from the different UI to the machine learning layer.

Machine Learning Layer: The machine learning layer is where plain text is recognized by NLP, NLU and NLG. The technology allows human natural language to be classified and extracted into intents and entities respectively.

* + 1. NLP: Natural Language Processing acts as a fundamental pillar for recognition of language. It makes the chatbot to understand the plain text and convert it into structured data.
    2. NLU: Natural Language Understanding is responsible for handling and converting formless data into a proper form that the chatbot can easily understand
    3. NLG: Natural Language Generation is responsible for the formation of linguistical correction of sentences and phrases

Data Layer: It consists of data connector and custom data source which are responsible for dataset training and API call in a chatbot system.

**2.3.1 Natural Language Understanding**

NLU and NLG are the subset of NLP. The principal duty is to make sure that Unstructured user communications are converted into intents and entities after which responds are returned to the user from the chatbot. The entire processes consist of the following

Tokenization: This is in charge of breaking string of words into discreate fragments known as tokens that bear a linguistic representation with a unique value for application

Normalization: Its duty is to analyse the text for any possible error or mistake such as misspelling or typographical errors that could change the supposed meaning of the users’ queries.

Entity Recognition: This process is in charge of recognising the entities by searching for a similar category of words, users’ date, or any other required information.

Dependency Parsing: This process in responsible for scrutinizing the text for nouns , verbs, subjects, objects, and phrases to find any dependent information conveyed by the user.

Sentiment analysis: This is in charge of forwarding the query to human being after assessing the users’ experiences and analysing the text by the chatbot software

**2.3.2 Data Processing**

The core functioning of chatbot depends on the amount of data trained. Organisations that are using chatbot must provide enough data. The data is the basis for managing conversations by chatabot following a question-answer pattern. The recognition of the question and the delivery of an appropriate answer is done by artificial intelligence and mechine learning. The figure 2.2 describes the Training data flow of each module in a separate channel. They are NLU which comprises of intent Classifier and Entity Extractor and Dialogue Policy which is taken care by Rasa Core Policy.

**“**Good morning”

You are welcome, I am covid Bot for Nigeria, how can I help you.?

Intent classifier

Custom entity extractor

Dialogue Policy

Training

Training

user

CovidBot

Figure 2.3: Data Training Flow

The user enters quarries in Natural Language, then the next is to tokenize and convert them into vector. They are some available techniques like kip-gram, word2vec and Continuous Bag of Word (CBW) to convert words to vectors followed by intent classification. The machine leaning Classification is handled by (Support Vector Machine) SVM or RNN-LSTM (Recurrent Neural Network – Long-Short-Term Model). This establishes the purpose of the intent of the user from the Natural Language. Another data process under NLU is entity extraction and Recognition which is done by Spacy-NER or LSTM or by using bidirectional LSTM, CNN and CRF. The dialogue police are taken care by the Rasa core

Entity Extractor

Entities

Text

Tokens

Features

Intent Classifier

Intents

Figure 2.4: The NLU Pipeline diagram

Tokenizers: This separates each word in the vocalization into a single token; typically, the tokenizer's output is a list of words. Tokenization is a crucial step in NLP jobs since it aids in preprocessing and representing text input in a way that is acceptable for additional analysis or modeling. In Rasa Framework, the following Tokenizer are integrated. 1. SpaCy Tokenizer, 2. Whitespace Tokenizer, 3. Custom Tokenizer. (Punitha Devi et al., 2021)

“You Are Welcome, I Am Covidbot for Nigeria, How May I Help You.?”

Tokenization of words

“You”, “Are”, “Welcome”,”,” “I”, “Am”, “Covidbot”, “For”, “Nigeria”,”, “How” “May”, “I”, “Help”, “You.?”

Figure 2.5: Tokenization by SpaCY Tokenizer

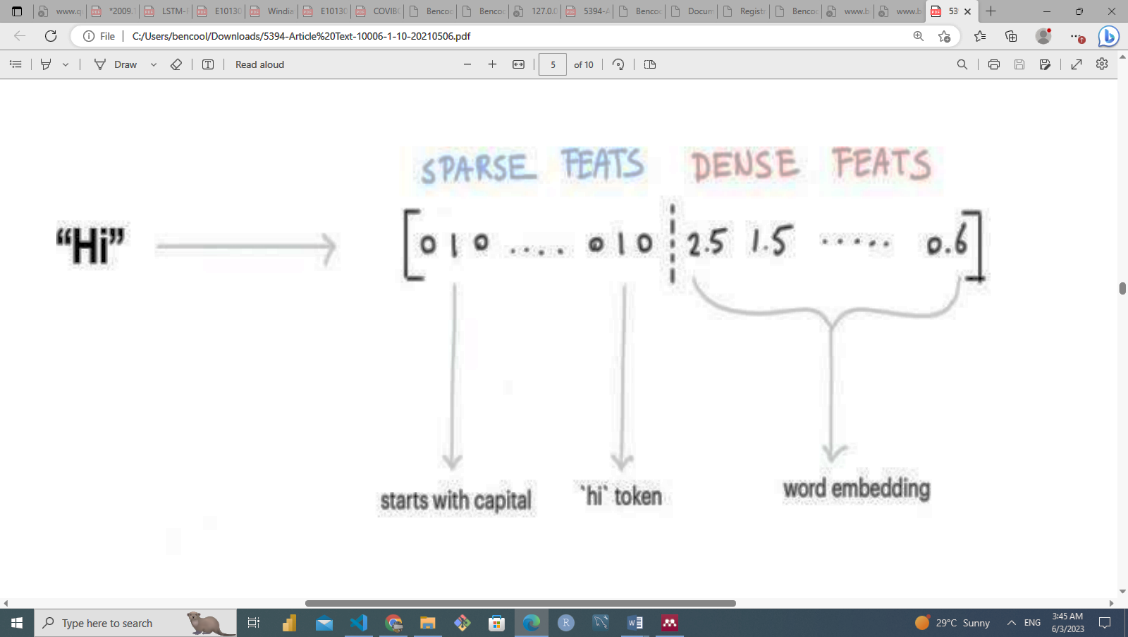
Featurizer:converts unstructured input data, such as user messages, into numerical representations that computer learning models can comprehend and exploit. Natural language understanding (NLU) and dialogue management duties heavily rely on featurisation.In the Rasa architecture, featurisation is a crucial step in training machine learning models for NLU and dialogue management tasks. It lets the models to comprehend the numerical representations of the input data and make predictions using those representations**.** For the purpose of converting text-based input into numerical features, Rasa has built-in featurisation algorithms. (Punitha Devi et al., 2021)

Figure 2.6 Featurizer (Encoding the word “Hi”)

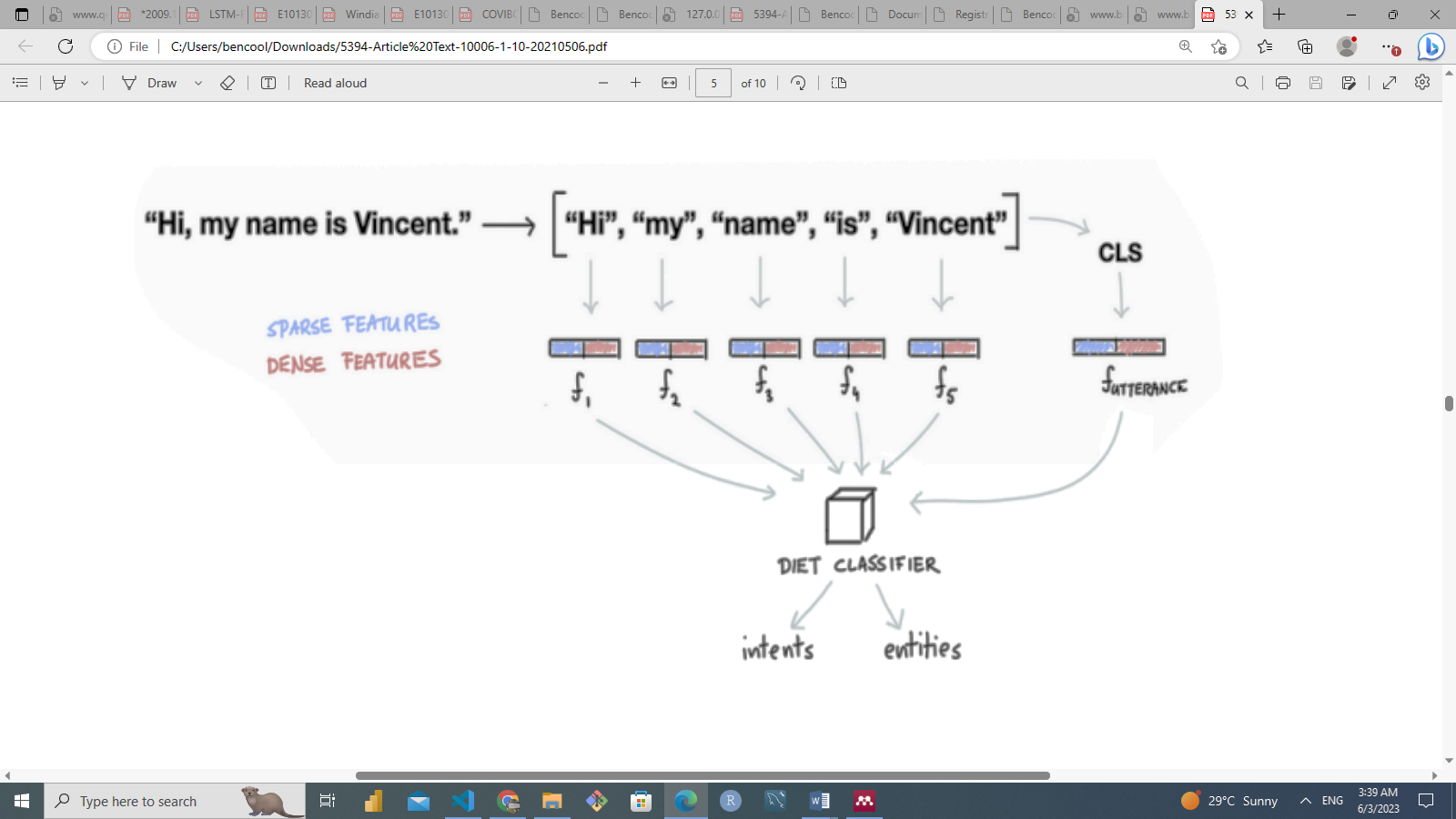
Intent Classifier: Rasa offers resources and elements for building intent classifiers, including pipelines for integrating custom models or rule-based classifiers as well as pre-built machine learning models. They are in charge of figuring out the meaning or objective of user input, such as text messages or spoken utterances. In conversation systems, chatbots, and virtual assistants, intent categorization is essential because it directs user requests to the proper actions or responses. In Rasa, Dual Intent Entity Transformer(DIET) model handles intent classification together with entity extraction (Punitha Devi et al., 2021)

Figure 2.7: Intent classifier **(**Punitha Devi et al.., 2021)

Entity Extractor: Identification and classification of named entities within text constitutes the core work of entity extraction, sometimes referred to as named entity recognition (NER), in natural language processing (NLP). Real-world items such as people, businesses, places, times, dates, time expressions, money values, and more are referred to as entities. Understanding the specific information specified in user input and extracting pertinent data for subsequent processing depend on entity extraction. Rasa offers built-in entity extractors, including extractors that use rules, like the CRF Entity Extractor, and extractors that use machine learning, like the DIET Classifier. Rasa also enables the use of pre-trained models and the integration of customized entity extractors for entity extraction.

Intents:An intent in the Rasa framework represents the objective or reason for a user's input. It's a technique for classifying and comprehending the user's aim, or what they hope to accomplish with their message. In a dialogue system, intentions assist in translating user inputs into the proper actions or reactions**.** An intent is a label or category that designates a particular user objective(Punitha Devi et al., 2021). The following are a few examples of intents: "greet," "book appointment," "get doctor," or "cancel appointment." The common practice is to correlate each user communication with a single intent, which represents the user's current intention.

Entities:Entities represent particular bits of data or input parameters in the context of the Rasa framework. Entities give the user's request or assertion more context and specifics. They are taken from the user's message and used to compile pertinent data to carry out the user's objective. Several entities can be extracted from the user input, for instance, “Book an appointment with a doctor at 12 pm at Federal Hospital Abuja on June 15th”

‘physician’: The entity representing the person to have an appointment with. Which is a doctor.

‘Location’: The entity representing the place to have the appointment which is Federal Hospital Abuja.

‘Time’: The entity representing the actual time of the appointment which is 12pm.

‘Date’ The entity representing the specific date of the appointment which is June 15th.

**2.3.3. RASA Stack's NLP Models Applications**

The Rasa stack is a flexible platform for creating conversational AI applications and consists of the Rasa NLU and Rasa Core parts. It can be used in a variety of practical situations where NLP is necessary. Researchers have unveiled a lot of instances which include the following:

Virtual Assistants: Rasa can be used to create virtual assistants for jobs like organizing appointments, providing customer service, or finding information. Users can communicate with the assistant using natural language inquiries, and Rasa can decipher their intentions and pullout pertinent entities to deliver the right answers.

Agri Farm Assistant for Farmer was developed with Rasa which took over the duties of call center to provide information to the farmers in India (Darapaneni et al., n.d.). Immediately the agricultural sector of India became super in the sense that all the shortcoming was conquered by the virtual assistance by providing and guiding farmers on issues relating to agriculture without delay unlike Agricultural experts.

Customer Support Automation: Rasa can automate customer care procedures by comprehending and allocating customer inquiries. It can categorize the customer's message's intent and extract pertinent data, enabling the system to give appropriate answers or escalate the request to a human agent as needed.(Punitha Devi et al.,2021) developed a banking chatbot known as B-bot that has been proven to be the efficient way to resolve customers queries.

## **2.4 Research Gaps Analysis**

## A clear comparison was carried out during the literature review between the existing chatbots systems and the present research study in order to visualize the efficiency of the system under the study. Following were the key consideration:

## Acovid-19 AI chatbot developed for Nigeria with a cloud-based framework known as Google DialogueFlow lacks the information regarding the flexibility of the chatbot codes to be customized and updated due to the limited covid-19 dataset and API connection in Nigeria. The study considered this as methodology gap, looking for the suitable framework that can accommodate the available dataset and upgrade it in future.

## A similar Covid-19 AI chatbot customized for another country which have failed to supply appropriate information for someone located in Nigeria. The study considered this as population gap, which looks for an appropriate field and label to be supplied to the user in Nigeria as a group or location

## Therefore, to fill in the identified gaps, this study considered the following solutions

## A non-cloud-based opensource AI framework known as Rasa will be used to develop a chatbot for the purpose of code flexibility

## A dedicated covid-19 chatbot will be develop for Nigeria. This is expected to be to capture all the environments and locations for easy tracking of the covid-19 patients.

**CHAPTER THREE**

**RESEARCH METHODOLOGY**

**3.1** **Research Design**

This section focusses on the research designs and techniques that were adopted in this study with the aim of achieving the research objectives. The goal of this study is to develop a Chatbot that will be capable of handling covid-19 Frequently Ask Quenstion together with emergency service referral functionality to profession doctors and nurses in Nigeria using Python and Rasa framework version 3.1. This will help us to tackle the population and methodology gap in the previous works. Lei et al.(2021) developed chatbot with Rasa Framework for American and Dada(2021) developed another one with Dialoguflow, a Google Cloud framework for Nigeria respectively.

Although, other multiple fundamental concepts and approaches that can be used for the development of a Chabot exists such as Watson Conversation (IBM Watson), Google Dialogflow and Language Understanding Intelligent Service (LUIS) but this study will adopt a branch of Artificial Intelligence and Deep Learning known as Natural Language Processing and Natural Language Understanding. For the purpose of achieving the set objectives effectively, Rasa framework version 3.1 was considered. This is an Open Source framework that allow developers to configure, deploy and run NLU on local server thereby given programmers full control on their codes for future modifications (Alotaibi et al., 2020).

**3.1.2 Rasa Framework**

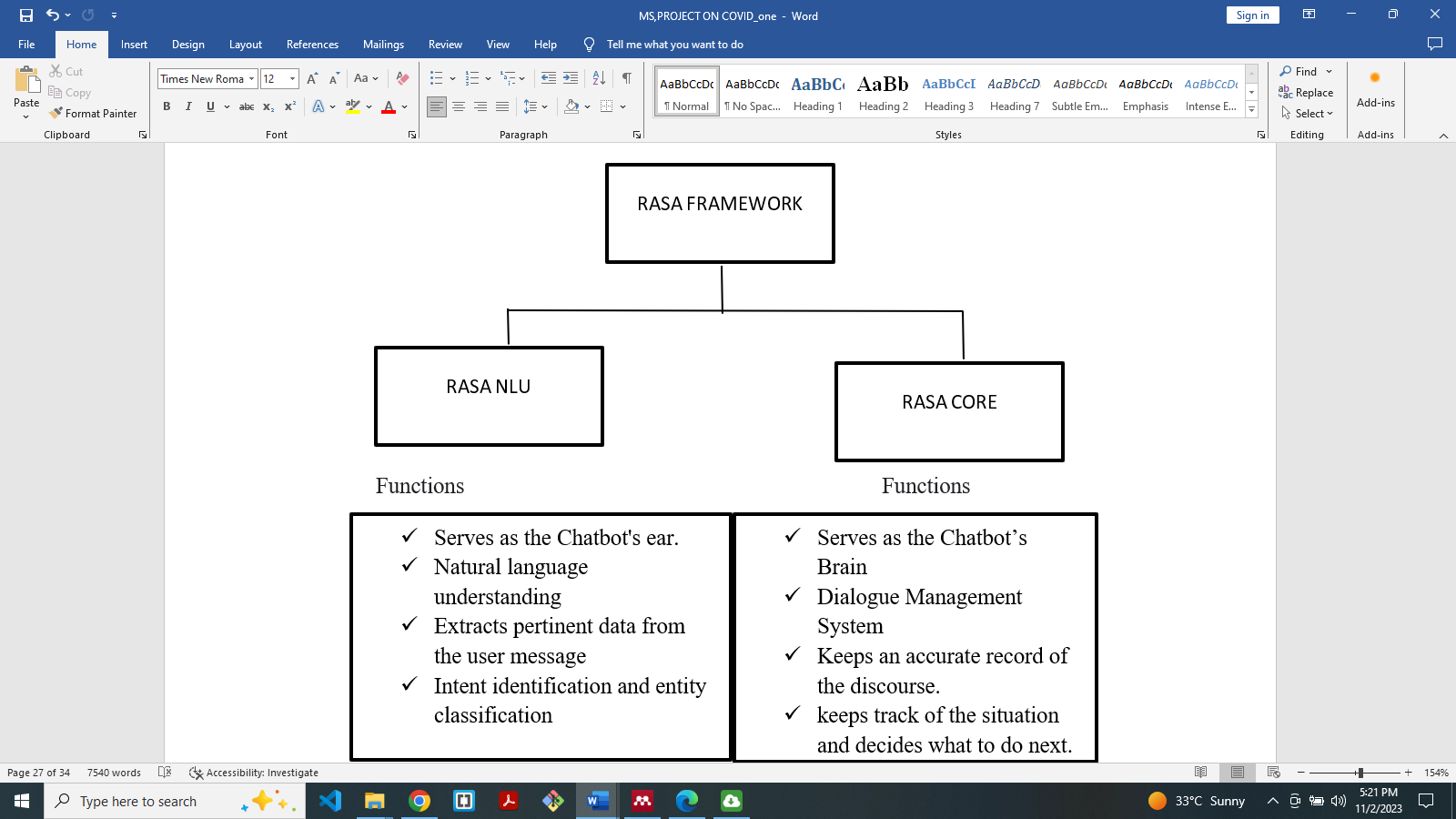
Rasa framework is a machine-learning framework that consists of two main components. First is Rasa NLU (Natural language understanding) and second is RasaCor**e** (Dialogue management) model. The Rasa NLU is like an ear which accepts the input from the user and Rasa Core (Dialogue Management) is like the brain which makes decisions based on the user input.

Figure 3.1 Overview of Rasa Components and Functions

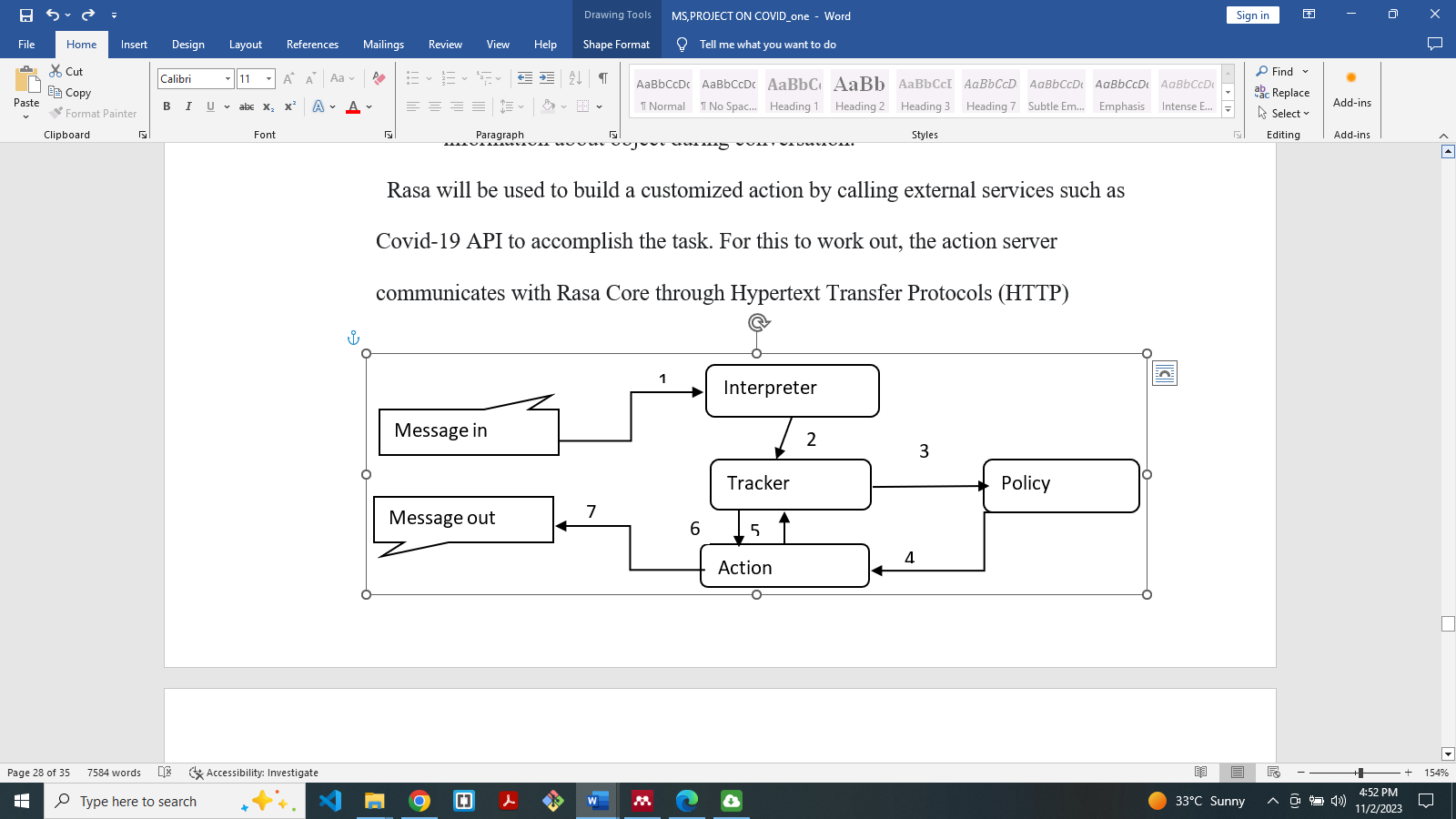
NLU Model is an algorithm that helps us to train our assistants to learn the meaning of words and the intention behind them. The major job of NLU is

1. Training data format: This is an input format for dataset used to impact a machine learning model.
2. language support: NLU has the ability to support any language
3. Choosing a pipeline: The choose of pipeline depends on the NLU
4. Entity Extraction: NLU determines the entity extraction method.

COREdecides the next action based on the current and historical dialogue records. It takes the structure input from the NLU and builds a probability model which decides the set of actions to perform based on the previous set of user inputs. The following will be under the control of Rasa Core

1. Dialogue Engine: This is the part of RASA open-source library that decides what to do next based on the context of the conversation.
2. Stories: They are the component of RASA library that is in charge of conversational flow.
3. Domain: This defines the space in which the chatbot functions by stipulating the intent, entities, slots, responses, forms and action that the bot should know about.
4. Responses: These are the messages which can be inform of texts and images that the chatbot sends to the user.
5. Action: These are the bot’s response to the user’s questions or messages
6. Forms: These are the conversation pattern that collects pieces of information from the user in order to determine the best response from the chatbot.
7. Knowledge Based Actions: These enable the user to obtain a detailed information about object during conversation.

Rasa will be used to build a customized action by calling external services such as Covid-19 API to accomplish the task. For this to work out, the action server communicates with Rasa Core through Hypertext Transfer Protocols (HTTP)

 Figure 3.2: Conversation Flow Diagram Of Covid-19 AI-Based Chatbot

The steps from the Figure 3.2 are

1. The interpreter receives the message that is passed to it and converts it into a dictionary including the original text, the intent, and any entity that was found, which is handled by NLU
2. The Tracker is the object which keeps track of the conversation state. It receives the into that a new message has come in
3. The policy receives the current state of the tracker
4. The policy chooses which action to take next.
5. The chosen action is logged by the tracker
6. The tracker returns it to action
7. A response is sent to the user

**3.2 System Design**

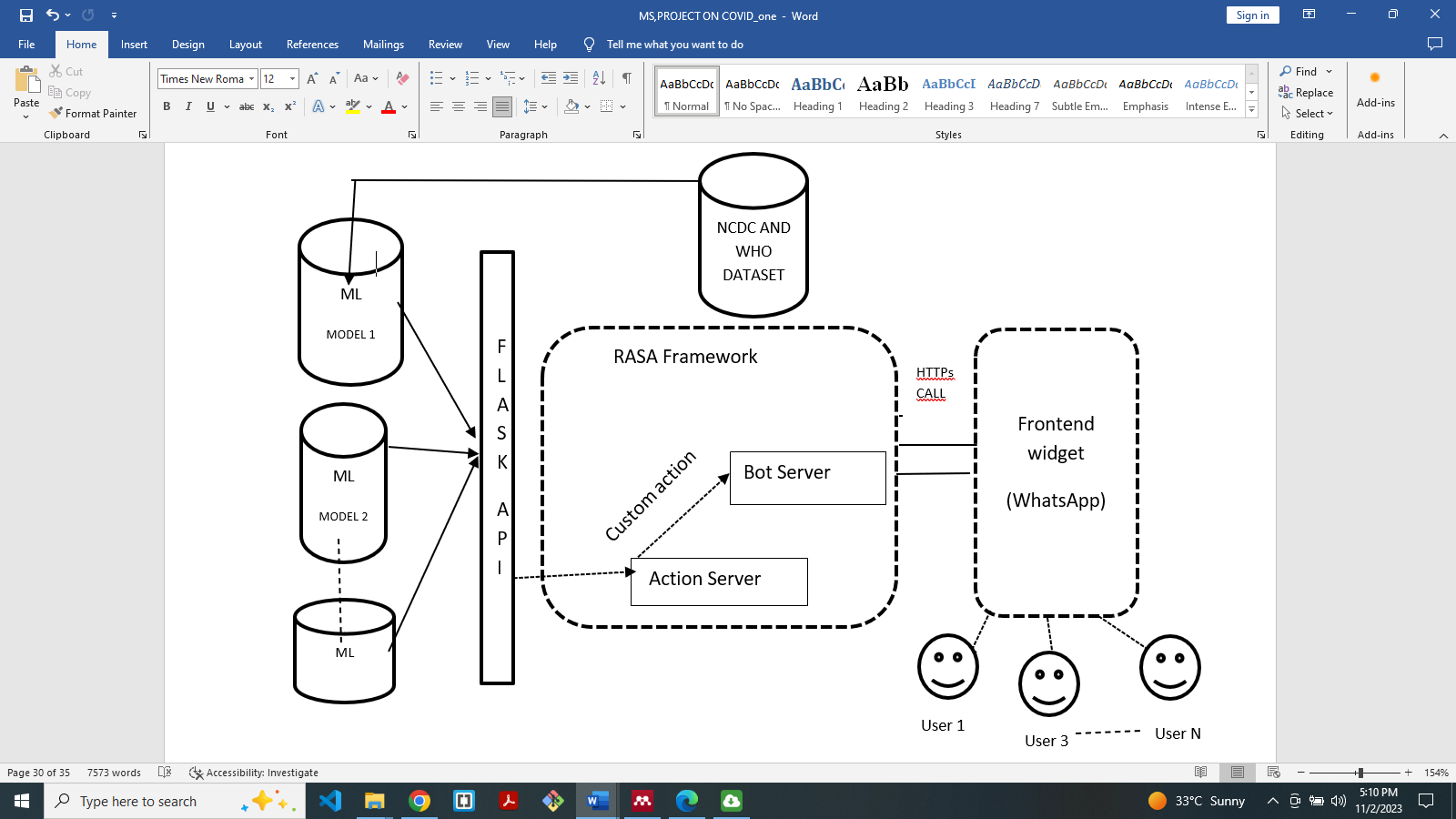
This study will adopt Structures System Analysis and Design Methodology (SSADM) to present COVID-19 AI BASED CHATBOT also known as N-CovidBot which is capable of resolving all the related Frequency Asked Questions (FAQ) on issues of coronavirus in Nigeria. The Chatbot’s model can be divided into three sections: Backend, Machine learning model and Frontend. The main Functionality of the chatbot is carried out by the Rasa Framework

Figure 3.3:Architectural Diagram of the Proposed Chatbot (N-CovidBot)

The figure 3.3 above describes the proposed chatbot system. The first component is the user who will query the system through questions or enquiries. The user does that with the help of the UI presented by the frontend widget. HTTPS CALL takes it to Rasa frame work which has other components such as action server, bot server and custom server. The action server connects to FLASK API (python framework) which will then take the data to be trained to the models (DIET and other machine leaning model pipeline) and then return them back to chatbot to display the output to the users

**3.6 Method of Data Collection**

This research work relies both on the secondary and primary data source. Dataset will be collected from reliable sources and certified website that have information about covid-19. Among sites are World Health Organisation website

(https://www.who.int/emergencies/diseases) and Nigeria Center for Disease Control (https://ncdc.gov.ng/themes/comm).

The collected dataset will be scrutinized by sampling opinions from medical personnels and individuals who have suffered and recovered from the disease across the country Nigeria through google form. From the collected data, Intents, Entities and Action data is expected to feed the system model for training.

**3.7 Technique of Data Analysis**

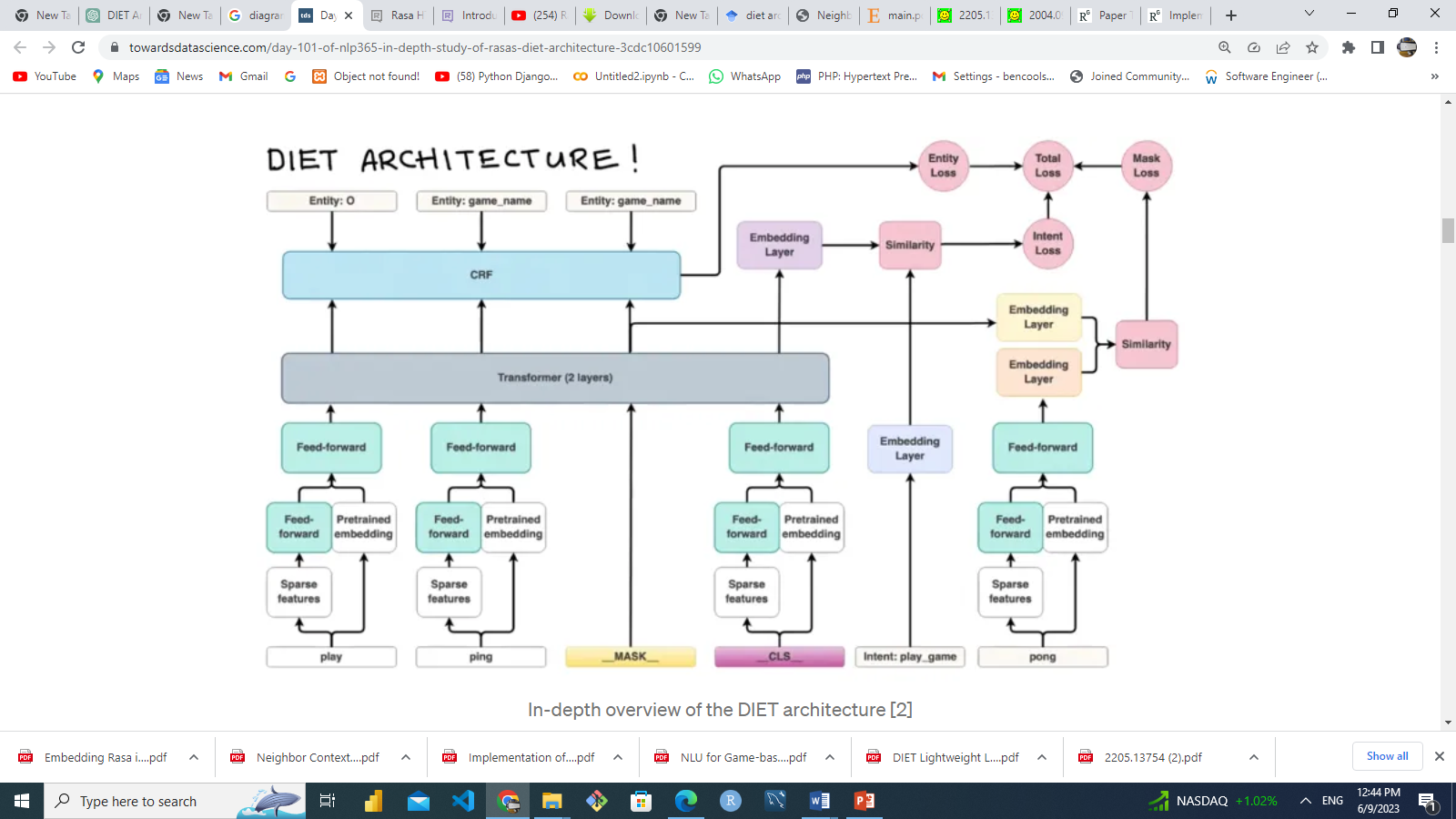
Most of the research in the field of machine learning is mainly quantitative since it involves the modeling of data with the existing data (supervise learning). Therefore, this study analysed a dataset collected from the relevant sources and feed them into the Rasa framework. Rasa framework has inbuilt techniques for data analysis and processing which include DIET. A DIET has the flexibility to swap out different components because it is made up of many separate parts (Wibawa et al., 2021). Its ability to integrate several word embeddings, Like BERT and GloVe, or trained words from the model language and sparse words. Also included in the n-gram character level in plug-and-play mode which is one of its key features that made it preferable for this study. It is a modular architecture that gives software developers more flexibility in their experiments (Bunk & Nichol, n.d.)

Figure 3.4 :The overview of DIET Architecture Adopted from (Bunk & Nichol, n.d.)

Figure3.4 shows how DIET architecture handles input data from the chatbot user. The request is to play a game know as ping pong. From the base of the diagram, you can see that the sentence is separated into separate words known as token (play, ping, pong). DIET then placed play game as intent and ping pong as entity (Entity: 0, Entity: play\_ game). From the play, \_CLS\_, and pong token, there are two directional movement of the tokens, one path to pretrained embedding algorithm (BERT, GLOVE, CoverRT) that will give a numeric vector as the output and the second path to sparse feature. The work of sparse feature it to generate one hot encoding and character gram of the token. The character gram will be concatenated and move together to Feed-Forward layer. DIET will apply an activation function which will handle all the vectors and losses from different tokens and give out the result.

The DIET architecture is preferred in the study because of its multi-task ability to handle both intent classification and entity extraction together. It provides the ability to plug and play various pre-trained embeddings (Bunk & Nichol, n.d.). Bunk, (2020) further confirmed that it is a modular architecture that fits into a typical software development workflow and parallels large-scale pre-trained language models in accuracy and performance.

3. 8 Justification of Methods

This study adopting an open-source AI framework conversational chatbot known as Rasa. Rasa 3.x will be used to conquer the weakness of the existing system for it has an inbuilt machine learning algorithm and models for intent classification and entity extraction at a very fast rate. The flexibility of the framework has given it more credit among the programmers who often develop chatbot for different function. It has capability to deploy and integrate other software and platforms.

In case of evaluation, Rasa has a robust command to present performance of the model on the intent, entity and confidence level. Others evaluations include confusion matrix, precision, accuracy, recall and f1-score.

**REFERENCES**

Almalki, M., & Azeez, F. (2020). Health chatbots for fighting COVID-19: a scoping review. *Acta Informatica Medica*, *28*(4), 241.

Baig, M. M., Meshram, A., Bansod, A., Mishra, N., Bagde, K., & Lad, M. A REVIEW ON SMART HEALTHCARE USING MEDICAL CHATBOTS.

Bunk, T., Varshneya, D., Vlasov, V., & Nichol, A. (2020). Diet: Lightweight language understanding for dialogue systems. arXiv preprint arXiv:2004.09936.

Herriman, M., Meer, E., Rosin, R., Lee, V., Washington, V., & Volpp, K. G. (2020). Asked and answered: Building a chatbot to address covid-19-related concerns. *NEJM Catalyst Innovations in Care Delivery*, *1*(3).

Jiao, A. (2020, March). An intelligent chatbot system based on entity extraction using RASA NLU and neural network. In *Journal of physics: conference series* (Vol. 1487, No. 1, p. 012014). IOP Publishing.

Omaka-Amari, L. N., Aleke, C. O., Obande-Ogbuinya, N. E., Ngwakwe, P. C., Nwankwo, O., & Afoke, E. N. (2020). Coronavirus (COVID-19) pandemic in Nigeria: Preventive and control challenges within the first two months of outbreak. African journal of reproductive health, 24(2), 87-97.

Patil, K., Shelke, J., & Raut, C. M. (2021). AI Based COVID-19 Awareness Chatbot.

Pokharel, B., Ganesh, K., Timilsina, B., Pokharel, Y., Makam, M., & Kaur, V. (2022, September). An Interactive AI-Powered Web Healthcare System. In *2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA)* (pp. 1-5). IEEE.

Prianto, C., & Harani, N. H. (2021). The Covid-19 chatbot application using a natural language processing approach. *IJISTECH (International Journal of Information System and Technology)*, *5*(2), 198-206.

Sharma, R. K., & Joshi, M. (2020). An analytical study and review of open source chatbot framework, rasa. *Int. J. Eng. Res*, *9*(06), 1011-1014.

Zhou, S., Silvasstar, J., Clark, C., Salyers, A. J., Chavez, C., & Bull, S. S. (2023). An artificially intelligent, natural language processing chatbot designed to promote COVID-19 vaccination: A proof-of-concept pilot study. *Digital Health*, *9*, 20552076231155679.