REPORT

1. Problem Definition

Children’s language development and cognitive skills play a crucial role in their overall growth and academic success.Traditional toys and learning tools often lack the ability to engage children in interactive conversations and provide personalized learning experiences.Children today are spending more time indoors and in front of screens, which can limit their opportunities for physical activity, socialization, and creative play. Nowadays parents are also looking for ways to supplement their children's education and provide engaging and interactive learning experiences Therefore,there is a need for an NLP speaking toy that can effectively understand children’s speech,respond appropriately,and facilitate language learning in an engaging manner.The problem at hand is to design and develop an NLP speaking toy for children that utilizes advanced speech recognition and natural language understanding techniques.The toy should offer a wide range of interactive activities and educational content to enhance children’s language skills and cognitive development.

2. Purpose and Need

(i) Enhancing Language Skills

One of the primary purposes of an NLP speaking toy is to promote language development in children. By engaging in interactive conversations with the toy, children can improve their vocabulary, grammar,sentence structure, pronunciation, and overall language proficiency.

(ii) Cognitive Development

The development of language is closely linked to cognitive abilities such as problem-solving, critical thinking, memory, and attention span. An NLP speaking toy can stimulate cognitive development by presenting children with challenging language tasks, puzzles, and educational content that require them to think, analyze, and respond appropriately.

(iii) Interactive Learning Experience

Traditional toys and learning tools often lack the ability to engage children in interactive conversations. An NLP speaking toy fills this gap by providing a dynamic and interactive learning experience. It can respond to children's queries, engage in dialogues, tell stories, ask questions, and provide feedback, making the learning process more engaging, stimulating, and enjoyable for children.

(iv) Social Interaction

Interacting with an NLP speaking toy can help children improve their social skills and emotional intelligence. Through conversations and role-playing, the toy can encourage communication skills.

3. Project Objective

- Develop Speech Recognition Capability:
  Design and implement robust speech recognition algorithms and models to enable the NLP speaking toy to accurately recognize and understand children's spoken words and phrases.
- Implement Natural Language Understanding:
  Develop advanced natural language understanding techniques and algorithms to enable the NLP speaking toy to comprehend the meaning, intent, and context behind children's queries, commands, and conversations.
- Design Interactive and Educational Content:
  Develop a wide range of interactive activities, educational content such as storytelling, quizzes, vocabulary building exercises, and language learning exercises to engage children and facilitate their language learning and cognitive development.
- Ensure Safety and Child-Friendly Design:
  Implement safety features, adhere to strict safety standards, and ensure the toy has child-friendly design elements and appropriate content and interactions to create a secure and suitable environment for children.
- Ensure Durability and User-Friendliness:
  Creates an intuitive user interface that is easy for children to interact with independently, considering their age and cognitive abilities.
- Evaluate Effectiveness and User Experience:
  Conduct thorough user testing and evaluation to assess the effectiveness of the NLP speaking toy in enhancing children's language skills, cognitive development, overall user experience and collect feedback .

4. Existing Methods / Models

(i) Voice Recognition-Based Toys

These toys primarily rely on voice recognition technology to understand

and respond to user inputs.They typically use pre-programmed phrases or patterns to recognize specific commands or trigger specific actions.However,their responses may be limited and less adaptable to natural language variations.

(ii) Rule-Based Systems

Some NLP speaking toys employ rule-based systems, where predefined rules and patterns are used to understand user inputs and generate responses. These systems often require extensive manual programming and lack the flexibility to handle complex conversations or understand nuanced language.

(iii) Custom NLU Engines

Some NLP speaking toys use custom-built natural language understanding (NLU) engines that are trained on specific datasets and models.These engines leverage machine learning techniques to understand user inputs and extract intents and entities. While they provide more flexibility than rule-based systems, they require significant development effort and ongoing maintenance.

5. Proposed Method / Model

The NLP speaking toy leverages Dialogflow, a powerful cloud-based NLP platform, to understand and interpret user inputs. Dialogflow employs advanced algorithms for intent recognition and entity extraction, allowing the toy to comprehend the user's commands and queries effectively. By utilizing Dialogflow's pre-built integration with the Angular web framework, the toy gains access to a robust development ecosystem and a user-friendly interface for designing conversational flows.The Angular web framework provides the infrastructure for developing the toy's user interface, business logic, and integration with external APIs. Through the Angular app, the toy processes the intent and entity information received from Dialogflow's NLU engine, generating appropriate responses or triggering specific actions. This combination enables the toy to offer a rich and customizable user experience.

It is better than the existing models because :-

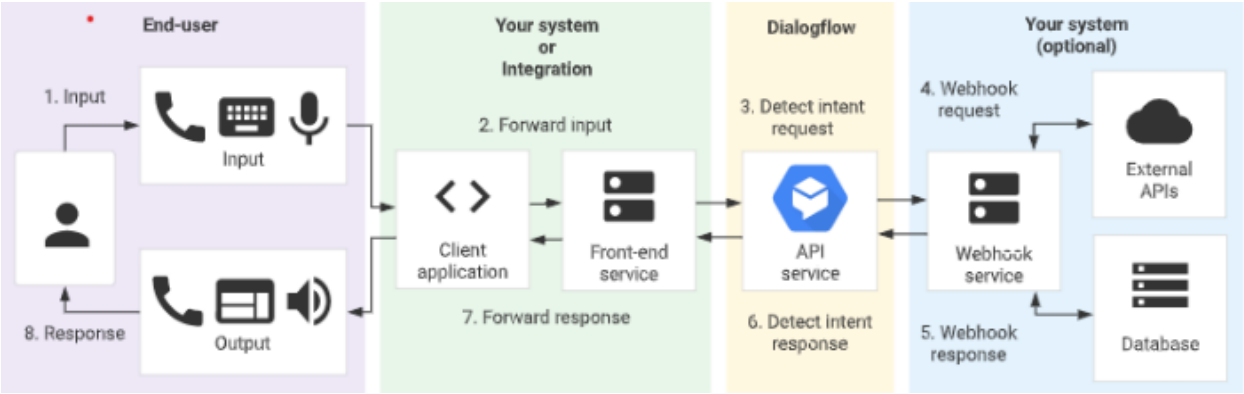
Advanced NLU Capabilities

Dialogflow is built on Google's powerful NLU technology, which has been developed and refined over years.Pretrained Models and Knowledge: Dialogflow comes with pretrained models and knowledge bases, covering a wide range of domains and common conversational scenarios.

Integration and Ecosystem

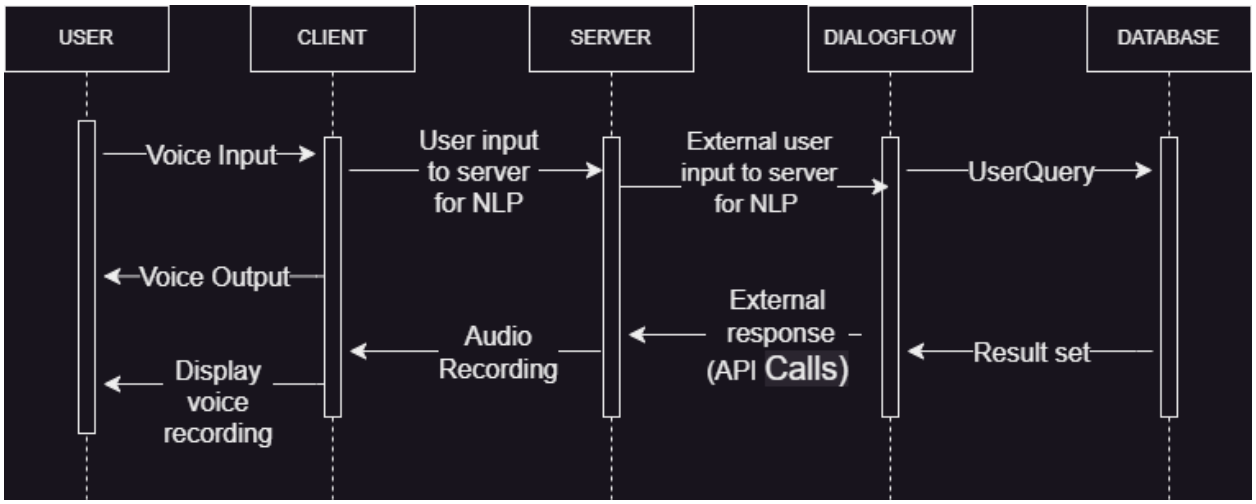
Dialog Flow offers seamless integration with various platforms and services, including web frameworks like Angular. It provides webhooks, APIs, and extensive documentation for easy integration and customization.Dialogflow is continuously

updated and improved by Google, ensuring that the NLP speaking toy benefits from the latest advancements in NLU technology.



6. Overall Architecture Diagram

7. Sequence Diagram or Use case diagram (or any UML diagram)



8. Module division

a) Client module

The Client module is responsible for receiving and sending audio data to and from the user's device. It handles the voice input and output functionalities. The user's voice is captured using a microphone or voice recording mechanism. It sends the recorded voice input to the server for processing and receives the generated responses. The generated voice response is then played back to the user and displayed on the screen.

b) Server module

The Server module acts as the central processing unit of the voicebot system. It receives the voice input data from the Client module, processes it using the Dialog Flow module, and generates appropriate responses. API calls are made for Cloud Speech to Text and the input extracted is forwarded to the dialogflow. Once the dialogflow processes input the response is extracted from it. API calls for Cloud Text to Speech are made to generate the audio output which is then send to the client module.

c) Dialogflow module

The Dialog Flow module is responsible for processing the natural language input from the user, understanding the intent, and generating appropriate responses. It involves natural language understanding (NLU) and natural language generation (NLG) components. Some preprocessing steps may be performed to clean and normalize the text (involve removing punctuation, converting text to lowercase,

handling abbreviations or acronyms, and dealing with any other specific requirements of your application). The preprocessed text is then passed to the natural language understanding (NLU) module which analyzes the text to determine the user's intent, extract relevant entities or keywords, and identify any specific actions or commands implied in the query.

9. Module wise diagram(Algorithm/diagrams)

1) Client module

Algorithm:

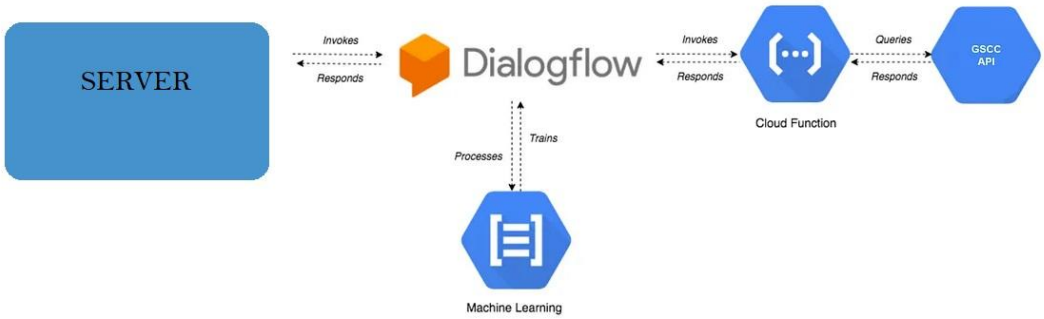
Step 1 : Start
Step 2 : Create pointers to the start & stop buttons.
Step 3 : Instantiate socket.io, and open a connection to the server.
Step 4 : Create 2 event listeners for starting and stopping the recording.
Step 5 : Capture the stream from the local microphone using navigator.getUserMedia() and configure the settings to optimize the stream using RecordRTC.
Step 6: RecordRTC API will return a string dataURL that contains your audio stream.
Step 7 : Create an object to send it to the server using socket.io
Step 8 : Stop

2) Server module

Algorithm:

Step 1 : Start
Step 2 : Importing all the required libraries
Step 3 : Loading the environment vars
Step 4 : Setting up the Express server with Socket.IO listeners
Step 5 : Google Cloud API Calls: Dialogflow Audio DetectIntent & DetectStream calls, Speech to Text Recognize & StreamingRecognize calls, Text to Speech synthesize calls
Step 6 : Stop

3) Dialogflow module



10. Assumptions

1. Only English language is allowed.
2. Language Proficiency: The toy assumes that the user has a basic level of language proficiency and can communicate in the language supported by the toy. It assumes that the user can articulate their requests or questions clearly enough for the toy to understand.
3. Reliable Network Connection: If the toy requires an internet connection for certain functionalities, it assumes that the user has access to a stable and reliable network connection. This is necessary for interacting with cloud services, retrieving data, or performing language processing tasks.
4. User Engagement: The toy assumes that the user will actively engage with it, providing input and responding to prompts or questions. It assumes that the user will actively participate in conversations and interact with the toy's features.
5. Clear Speech Input: The toy assumes that the user will provide clear and understandable speech input. It assumes that the user will speak in a way that the speech recognition system can accurately transcribe and process.
6. Privacy and Security: The toy assumes that the user values their privacy and expects the toy to handle their data securely. It assumes that the user is willing to provide certain personal information, such as name or preferences, for a personalized experience but expects that sensitive information will be protected.
7. Age Appropriateness: If the toy is designed for a specific age group, it assumes that the user falls within that age range and will interact with the toy accordingly. It assumes that the user will understand and comply with any safety guidelines or content restrictions specified for their age group.
8. Safe Usage: The toy assumes that the user will use it in a safe and appropriate manner. It assumes that the user will not misuse the toy or engage in behavior that may cause harm to themselves or others.
9. Proper Maintenance: The toy assumes that the user will perform basic maintenance tasks as recommended, such as replacing batteries, charging the

toy, and keeping it clean. It assumes that the user will follow any instructions provided for the toy's upkeep.

10. No Speech Impairments: The toy assumes that the user does not have significant speech impairments that could hinder communication. It assumes that the user's speech is clear and intelligible enough for the toy's speech recognition system to understand.
11. Limited Context Awareness: The toy assumes a limited context awareness, meaning that it may not remember every detail or context from previous interactions. It assumes that the user will provide enough information or context to make the conversation coherent but may not have an extensive memory of past conversations.

11. Work breakdown and Responsibilities:

Work Breakdown:

Roshni Joshy:

- User Interface (UI) Development
- Audio encoding
- Dialog Flow processing

Sneha Elsa Mathew:

- Speech recognition
- Intent recognition
- Connection to database
- Dialog Flow processing

Roze Susan Mathew:

- Entity extraction
- Dialog management
- Response generation
- Sending audio data
- Dialog Flow processing

Swathi P :

- Audio capture
- Audio playback
- Receiving audio data
- Dialog Flow processing

Responsibilities:

Roshni Joshy:UI Designer & Back-end Developer
Sneha Elsa Mathew: NLP Specialist & Back-end Developer
Roze Susan Mathew: NLP Specialist & Back-end Developer
Swathi P. : Front end & Back-end Developer

12. Hardware and Software Requirements for Development

Hardware Requirements

System Architecture :Intel Core i7 or equivalent
RAM:>=4GB
Hard Disk>=: 32GB

Software Requirements

OS: Windows 11
Technology:
Angular CLI: v12.0.0 or later
DialogFlow CX
Database : MongoDB
Frameworks of Node: Express
IDE: Visual Basic Studio

13. Module wise Scheduling (Mile stones with week wise plan):

Dialog flow module

- Speech recognition
- Intent recognition
- Entity extraction
- Dialog management
- Response generation

Client module

- Audio capture
- Audio playback
- Data transmission

Server module

- Receiving audio data
- Dialog Flow processing
-Connection to database
- Audio encoding
- Sending audio data

WEEK 1 : (1 June-8 June)

- Develop client side UI design
- Setting up audio capture
- Speech recognition & Intent recognition
- Entity extraction & Dialog management

WEEK 2 : (9 June - 15 June)

- Data transmission
- Audio playback
- Response generation

WEEK 3 :(15 June- 22 June)

- Audio encoding
- Sending audio data
- Receiving audio data
- Connection to database

WEEK 4 :

- Dialog Flow Integration

14. Risks and challenges

1. Speech Recognition Accuracy
   Achieving high accuracy in speech recognition can be challenging, especially for children's speech, which can vary significantly in pronunciation, articulation, and clarity. Ensuring robust and accurate speech recognition performance in various environments and with different accents and speech patterns poses a significant technical challenge.
2. Data Privacy and Security
   Ensuring compliance with privacy regulations and implementing robust security measures to protect children's data is critical to maintaining trust and meeting legal requirements.
3. Validation and Evaluation
   Assessing the effectiveness and impact of the NLP speaking toy on children's language skills, cognitive development, and educational outcomes requires rigorous validation and evaluation. Conducting appropriate user testing, collecting feedback, and measuring learning outcomes can be resource-intensive and time-consuming.
4. Ethical Considerations
   Developing an NLP speaking toy for children involves ethical considerations, such as ensuring appropriate use of data, avoiding bias in speech recognition and language processing, and promoting responsible and safe interactions. Addressing these ethical considerations requires careful attention and adherence to ethical guidelines and best practices.
5. Technical Limitations and Resource Constraints

Technical limitations, such as processing power, memory, and storage constraints, may impact the functionality and performance of the NLP speaking toy. Additionally, resource constraints, such as development time, budget, and availability of skilled personnel, can pose challenges to project execution.

15. Conclusion

The NLP speaking project utilizing Dialogflow and the Angular web framework is expected to demonstrate the ability to create an interactive, conversational, and user-friendly toy experience. The implementation of the project highlights the effectiveness of these technologies in developing NLP speaking toys with enhanced language understanding, customization options, and a dynamic user interface. Moving forward, further enhancements can be made to the project by incorporating additional functionalities, such as integrating with external IoT devices, expanding the range of supported commands, or refining the user interface for an even more engaging experience. The project lays a solid foundation for future developments in the field of NLP speaking toys, offering opportunities for innovation and exploration of new possibilities in interactive toy technology.

16. References

[1] M. Semilof. (1996, July). “Chatbot Using Dialogflow and Web Services”.[Online].Volume-9 Issue-1S.
Available:https://www.ijitee.org/wp-content/uploads/papers/v9i1S/A10521191S19.pdf
[2] L.Boonstra. (2021). “Building Your Own Conversational Voice AI With Dialogflow” [Online]. Available: https://www.leeboonstra.dev/chatbots/building-your-own-voice-ai-1/
[3] L.Boonstra. (2019). An Overview of Chatbot Technology[Online]. Available: https://www.leeboonstra.dev/chatbots/contest/
[4] M. Holland. (2002). “Building a Simple Chatbot” [Online].
Available:
https://www.smashingmagazine.com/2017/08/ai-chatbot-web-speech-api-node-js/