

## Potential Tools to Support Learning: OpenAI and Elevenlabs Integration

Roro Ayu Fasha Dewatri<sup>1</sup>, Achmad Zahin Al Aqthar<sup>1</sup>, Handi Pradana<sup>2</sup>, Betuah Anugerah<sup>2</sup>,  
Wakhid Herry Nurcahyo<sup>3</sup>

<sup>1</sup> Teknik Informatika, Universitas Negeri Surabaya, Surabaya

<sup>2</sup> Department IT & Knowledge Management, SEAMOLEC

<sup>3</sup> Teknik Komputer dan Jaringan, SMKN 1 Banyumas, Jawa Tengah

[roroayufasha16@gmail.com](mailto:roroayufasha16@gmail.com)

### Abstract

The development of technology in education has grown significantly, where technology is now a principal aspect in the communication, storage and transfer of information, as well as the production of audio-visual media. This article discusses the design and implementation of combining two innovative technologies, OpenAI and Elevenlabs, to create potential learning aids. OpenAI, using generative language models such as GPT-3, provides an opportunity to develop virtual assistants with realistic communication capabilities in a learning context. In contrast, Elevenlabs offers voice artificial intelligence technology that enables the creation of highly realistic human voices in multiple languages. The utilisation of Raspberry Pi as hardware provides flexibility in controlling the interaction between AI and physical devices. Third-party platforms from OpenAI and ElevenLabs enable realistic and effective voice interaction in educational environments. In this discussion, the article details the advantages and limitations of each technology, including the voice realism of ElevenLabs and the natural language processing capabilities of OpenAI. The integration of OpenAI and ElevenLabs is expected to enhance the learning experience, while considering the financial and technical challenges that may arise.

**Keywords :** Learning, Integration, Artificial Intelligence, ElevenLabs, OpenAI

### Background

Educational technology is undergoing significant changes due to the widespread use of computers (Tuma, 2021). Technology is an essential part of communication, information storage and transfer, and the use and production of audio-visual media. The role of technology has evolved from a mere tool of study and enquiry to an integrated approach and use in education (Character Education for the Digital Age, n.d.). In this context, this article will explore the challenges that arise in transforming text into natural sound, creating a conversation that can enrich the learning experience. In this research, we reveal the integration efforts between two innovative AI technologies, OpenAI and Elevenlabs. In an effort to find solutions to these challenges, this article will detail related literature and research. The literature review will illustrate an in-depth understanding of the concepts of artificial intelligence, natural language processing, and voice technology. In addition, supporting theories that support the integration of OpenAI and ElevenLabs will be presented. This theory includes a conceptual foundation for addressing the challenges of voice-to-text conversion and conversely. This theoretical understanding forms the basis for further research on how the integration of these two tools can create voice output that is similar to the human voice in a learning context.

OpenAI, as one of the tools in advanced development has launched various highly advanced generative language models, such as GPT-3 (Gilson et al., 2023). With the ability to understand and generate text that almost resembles human writing, OpenAI allows us to develop virtual assistants that are capable of having

conversations that feel real in various contexts, including learning. On the other hand, Elevenlabs is a company that focuses on voice AI research and development by creating the most realistic, versatile and context-aware audio AI. The technology can generate speech in hundreds of new and existing voices in more than 20 languages (ElevenLabs: AI Voice Generator & Text to Speech, n.d.). In addition, the project will utilise the Raspberry Pi device, an efficient and powerful minicomputer that has dimensions roughly equal to the size of a credit/debit card (a smaller version of a modern computer) that is capable of performing tasks effectively. It was created by the UK Raspberry Pi foundation in the hope of enlightening and empowering a generation of students to be more creative and efficient (Dipak Ghael et al., 2020). The Raspberry Pi provides a strong foundation to support the operation of the tool.

The two main focuses in the article are identification in the context of audio or voice. (1) Overcoming the challenges of converting voice to text and conversely. (2) The resulting voice will be human-like enabling a conversational experience similar to live communication with a real human. By synchronising these two aspects, the aim of this article is to enhance the development of the tool by integrating two artificial intelligences, namely OpenAI and Elevenlabs.

## Literature Review

The literature review in this article presents an analysis of various sources of information covering the concept of artificial intelligence, the role of Raspberry Pi, the use of the Python programming language, the exploration of OpenAI technology, and the contribution of Elevenlabs in the context of this research. In the context of the literature presented, the concept of artificial intelligence refers to a field of computer science that focuses on the development of algorithms and computational models that can give machines the ability to learn and adapt automatically. First things, Artificial Intelligence is a branch of knowledge in computer science known as AI which can make a machine (computer) run and do work as well as humans do (Raje et al., 2021). In another sense, Artificial Intelligence is a branch of computer science that represents knowledge using more symbols than numbers and processes information based on several rules (*Artificial Intelligence (AI) | Definition, Examples, Types, Applications, Companies, & Facts | Britannica*, n.d.).

Raspberry Pi or what can be called RasPi is a credit card-sized computer developed in the UK by the Raspberry Pi Foundation. Raspberry Pi is a single "on-chip" computing hardware (Durga Prasad et al., 2018). Due to its small size, RasPi can be carried everywhere. In addition, RasPi also requires very little electricity to operate, so Raspberry Pi commonly called RasPi will save costs in electricity usage. RasPi is equipped with built-in WiFi, 4 USB ports, 40 GPI pins), a full HDMI port, an Ethernet port, a combined 3.5 mm audio jack, a Camera Interface (CSI), microSD card slot. Raspberry Pi 3 has a LINUX-based operating system called Raspbian (Grimmett, n.d.). The R-pi board contains 40 general-purpose input-output (GPIO) pins that can be used for digital input and digital output, contains 4 USB ports, 1 HDMI port, a 3.5 mm Audio Socket, micro USB power supply. This board also has a serial connection to connect the camera and display (Abraham & Nithya, 2021; Vashistha et al., 2019).



Image 1 Raspberry Pi  
(Source: [www.amazon.com](http://www.amazon.com))

Python is a powerful general-purpose programming language designed by Guido van Rossum in 1989. It is classified as a high-level programming language because it automatically handles the most basic operations (such as memory management) performed at the processor ("machine code") level. It is considered a higher-level language than, for example, C, due to its expressive syntax (which is close to natural language in some cases) and rich variety of native data structures such as lists, tuples, sets, and dictionaries (Hill, 2016).

OpenAI is a non-profit organisation that aims to develop AI technology with one of its most popular products being ChatGPT. ChatGPT is a technology that allows users to interact with text-based machines. It uses the GPT-3 (Generative Pre-Trained Transformer) algorithm to understand the context and generate relevant responses (Gilson et al., 2023). Elevenlabs is an AI Voice Generator platform owned by New York City, allowing users to create a new synthetic voice like a person's voice. In another sense, Elevenlabs is an AI-driven online text-to-speech software whereby providing text input, the text will be converted into spoken audio and can be saved as an MP3 file (Chen, 2023). This text-to-speech technology enables synthetic, cloned, or true AI artificial voices that mimic different genders, ages, and ethnicities. This technology is also capable of providing the ability to generate speech in hundreds of new and existing voices in more than 20 languages (*ElevenLabs: AI Voice Generator & Text to Speech*, n.d.).

## Research Method

The research design methodology approach focuses on systematic and organised steps, it forms a framework for structuring and conducting research with the aim of gaining a deep understanding of the topic under study. It involves a series of steps such as designing, implementing, and testing prototypes or tools. The research design always determines the type of analyses that will be conducted to obtain the desired results. The research design articulates what data is needed, what methods will be used to collect and analyse the data, and how the data will answer the research questions (Asenahabi, 2019). So, it can be said that research design is a procedural plan adopted by a researcher to answer questions in a valid manner with utmost objectivity and accuracy (Schantz & Lindeman, 2019). Research design is defined as a framework of methods and techniques chosen by a researcher to combine the various components of a study in a logical enough manner so that the research problem can be efficiently addressed (Boru, 2018). Explanatory research designs are conducted to explore research topics in depth when the research area is new and not widely researched (Akhtar et al., 2016).

The main focus of this research is on the tool design that includes the integration of artificial intelligence technologies from OpenAI and ElevenLabs. The steps in problem identification, solution design, tool implementation, internal testing, and related evaluation and improvement will be presented in detail to provide a comprehensive insight into the approach taken. (1) Identify the problem or gap to be solved or filled in the research context described in the preface in the research objectives. Then, (2) designing a suitable solution. The initial design focuses on the integration of technologies from OpenAI and ElevenLabs. This integration concept will be outlined in the tool design, including the system architecture, user interface, and implemented AI modules. A detailed explanation will be given on how these elements are integrated to achieve the research objectives. (3) Implementation of the tool by building a prototype with a predefined design. The prototype implementation stage includes code development, functionality testing, and iteration to improve the quality of the tool and relevant technical specifications will cover the implementation aspects. This includes system requirements, programming languages used, technology infrastructure, and other technical parameters. The technical details will provide an understanding of how the tool was developed with a special focus on the integration of OpenAI and ElevenLabs which was tested through comparison using voice from Google. (4) Internal trials focusing on comparing the integration of OpenAI and Google with OpenAI and ElevenLabs. The results of this comparison will provide insight into the most suitable integration to use by considering the advantages and disadvantages of each. After getting the results from the internal trial, the last step is (5) evaluation and improvement. Evaluation is conducted on the performance of the tested integrations, and the results are used to make necessary adjustments or improvements to enhance the effectiveness of the developed tool.

## Implementation

The learning tool that will be created involves the integration of two artificial intelligence technologies, namely OpenAI and ElevenLabs. The implementation involves inserting these two technologies into a mini computer called Raspberry Pi. The system is designed to allow users to interact with the device using voice. The basic concept is that the user will provide voice input which will be captured and converted into text. The resulting text will then be used as input for the OpenAI API, which will provide a text response. The response is then converted back into voice (Carlini & Wagner, 2018). The next Elevenlabs technology is used to improve the sound quality produced by wireless external speakers. With the application of Elevenlabs' audio-to-text

processing technology, the sound produced will be clearer and of higher quality. The implementation of Elevenlabs in this project is when the software through voice synthesis converts text into voice. This process is where the text response is passed to Elevenlabs to generate the native speaker's voice. In addition, we used Elevenlabs in this project to clone the voice that can be played repeatedly. To give a clearer picture of how this system works, the following flowchart illustrates the workflow of the device. (Fig. 3).

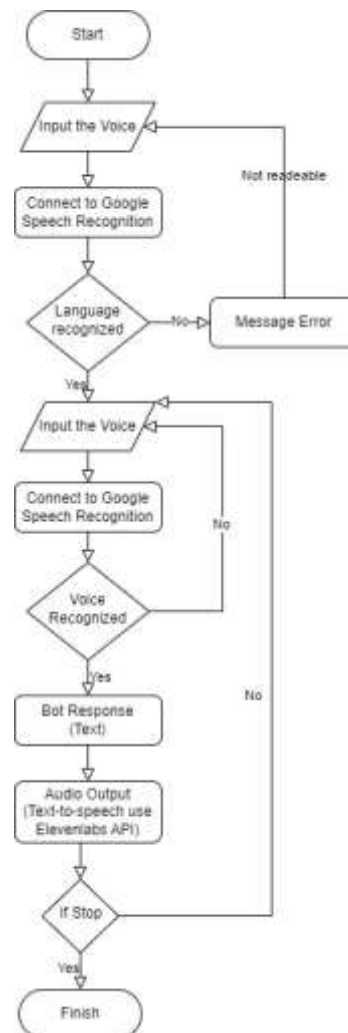


Image 2 Flowchart of the System

The flowchart of the system starts from the starting point (Start), where the system will provide a greeting to the user in the form of a voice. Then, the user will give voice commands to select the language they want to use on the system that supports English and Indonesian. The user's voice will be detected by the system using Google Speech Recognition. If the user's voice is not detected, the system will display an error message and ask the user to repeat the selection of the desired language. After the user's voice command is successfully detected, the system will process it and respond with the selected language. The user can proceed by giving a command or question to the system and start a conversation where the question will be sent to the OpenAI API for a response. The received response will be returned to the system in text form. However, if the user's voice is not detected when asking the question, the system will ask the user to repeat the question. The answer is then converted into voice using ElevenLabs' text-to-voice AI technology. If the user feels they are done with the conversation, they can end it by saying stop. The system will detect the user's voice and end the conversation if detected. If it is not detected, the system will ask for voice input again to continue the conversation.

Turning to the hardware used, the Raspberry Pi is the core of this system as a Linux-based mini computer. Raspberry Pi was chosen for its small size, low power consumption, and flexibility in integrating various additional hardware components. In addition, we used external hardware, including a microphone and wireless speakers connected to the Raspberry Pi via a USB port. We used a Raspberry Pi 3 model B. The Raspberry Pi acts as the main brain of the system by running software that supports voice recognition, and voice processing, and controls the voice output. Connection to the Raspberry Pi can be done through a Wi-Fi network that allows us to access it. For voice input, a wired external microphone with a USB interface is used. This microphone allows the capture of voice sounds with high sensitivity and quality suitable for this application. The advantage of a wireless microphone is flexibility in placement. As sound output, an external wireless speaker connected via the USB port was used. This speaker provides sound output. The use of wireless speakers allows for optimal placement so that the sound can be heard well by the user. The integration of the Raspberry Pi with the wireless microphone and speaker allows the system to receive voice commands from the user and respond effectively through voice. The hardware design scheme that will be used for the implementation of OpenAI and Elevenlabs will look like in Figure 4.

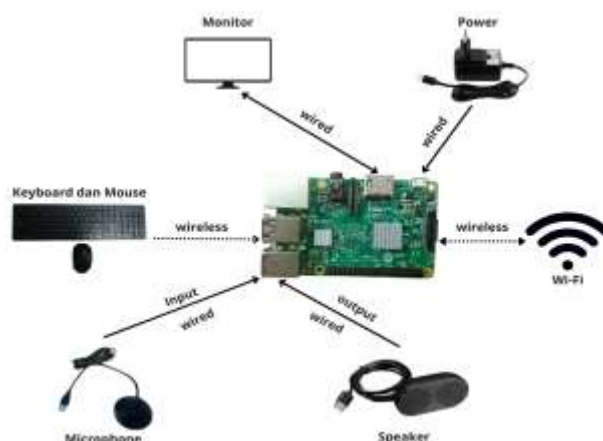


Image 3 Scheme of Hardware Design

The system is centred on the Raspberry Pi which serves as the core of the device. In development, we ensured that the Raspberry Pi is capable of managing voice interaction. A wireless external microphone is attached via a USB port, playing a role in collecting voice commands from the user. A wireless external speaker was also connected to deliver voice responses. To support natural language understanding and response, OpenAI services, specifically GPT-3, were integrated into the software. GPT-3 is used to process the received voice commands, generating relevant text responses. This text is then converted through speech synthesis. Implementation of OpenAI in the project: (1) The user speaks a command such as "Hello, How Are You?" (2) The microphone captures the audio and sends it to the Raspberry Pi. (3) The software will convert the voice into text using Google Speech Recognition. (4) The text will be forwarded to openAI's GPT-3 service for processing. (5) GPT-3 returns the text containing the greeting and the question "Can I help you?" (7) Elevenlabs through speech synthesis converts the text into voice. (8) An external speaker sends the voice to the user.

## Program Testing

This section will discuss testing the programme of the tool that has previously been designed in the form of a prototype scheme. This programme will be inserted into the Raspberry Pi so that it can be run. Testing will be done manually via a command prompt and using a user interface made in the form of a website display (user interface). The initial stage of this project is connecting the Raspberry Pi to a WiFi network. We do this to ensure that RasPi can connect to the internet because this project requires an internet network to run. The initial configuration process begins by activating the auto-connect option on the Raspberry Pi. Thus, the device will automatically connect to a predetermined WiFi network every time RasPi is switched on. Through this auto-connect method, RasPi will automatically search for and connect to the specified WiFi network. After successfully

configuring auto-connect, RasPi is ready to continue the development project. Our success in connecting to WiFi on RasPi can be seen in Figure 5 below..

```
pi@raspi:~$ cd projectAI/
pi@raspi:~/projectAI$ sudo python nm_connect.py
WiFi is enabled
Your current IP Address: 192.168.50.234
Device 'wlan0' successfully activated with '76d8e904-f985-41ad-a51c-4539cd7d6a89'
Koneksi Wi-Fi telah diubah.
Your IP Address: 192.168.40.116
pi@raspi:~/projectAI$
```

Image 4. Display of Connect WiFi on RasPi

The third party we used in this research is OpenAI, a powerful platform that enables natural language processing and provides relevant responses through access to its API. The natural language response provided by OpenAI enhances voice interaction through conversational access that provides quite varied responses. Each time a user uses it by giving a command or question, OpenAI will give a different response even with the same question. The use of OpenAI in this project provides effectiveness in the context of voice-based learning..



Image 5. OpenAI Display, Chat GPT 3

However, it should be noted that one of the main limitations of using OpenAI is the financial aspect. OpenAI is a paid service which means users have to pay to access the features and services provided. Other limitations faced when using OpenAI include; limited validity: The data available in OpenAI is quite limited as the database only covers information up to 2021. Reliance on Internet connectivity: Using OpenAI requires stable Internet access, which can be a problem in environments with poor connectivity.

The voice quality results produced by ElevenLabs' technology show significant success, i.e. a voice similar to a real human voice, with the default voice in English. ElevenLabs also provides voice technology in multiple languages, including Bahasa Indonesia. Thus, the same voice can be used in multiple languages. ElevenLabs is a text-to-speech application that provides the following benefits.

(1) Realistic Text-to-Speech technology: ElevenLabs delivers native English-like voices by default, providing a very realistic audio experience. (2) High-quality Audio Output: Clear and easy-to-understand voice quality that is effective in voice interaction. (3) Voice Cloning Capability: The ability to clone a voice by modelling an existing voice within the multiple voice options in ElevenLabs.

Disadvantages of ElevenLabs are (1) Paid: It is a paid service with a character limit on free usage of 10,000 characters. This means users have to pay a fee to access the features and services provided by ElevenLabs. (2) Technical Limitations: Limitations in generating voice variations that reflect emotions while speaking.



Image 6. ElevenLabs Display

In testing the programme in voice-to-voice mode, the user was first given the option to select a language, either English or Indonesian (Figure 8). After selecting the language, the conversation started. During the conversation, the user can ask questions or give commands using voice and the programme automatically responds with the pre-selected language (Figure 9). In addition, the user has the flexibility to spell abbreviated words or acronyms such as 'SEAMOLEC' and the programme will understand the meaning of the words based on the context of the conversation (Figure 10).

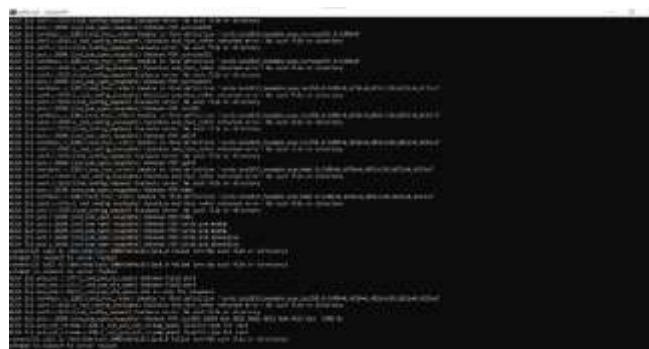


Image 7. Start Running Program Display



Image 8. Conversation with AI



Image 9. Word Recognition Abbreviations

In addition, users are given the flexibility to switch languages in the middle of a conversation by giving the 'switch' command. The system will ask again for the command to select English or Indonesian. After that, the user can start the conversation as when selecting English before. The following image shows an Indonesian conversation with the machine.



Image 10. Conversation with AI using the Indonesian Language

The use of OpenAI and ElevenLabs integration in this research has opened the door for robust voice and text integration in AI-based conversation models. The integration process starts with converting voice into text, which is then channelled to OpenAI via API to get AI-based text responses. The response from OpenAI, in the form of a relevant answer, is converted back into a native English or Indonesian voice using ElevenLabs technology. A significant contribution of this approach is to create interactive learning, where the AI conversational model becomes the bridge that connects the user's voice with informative and engaging responses. Using Bahasa Indonesia with the same voice allows for more personalised interactions in an educational context, thus helping to improve the quality of learning and user engagement.

## Result and Discussion

In this section, a study of voice perception in the development of AI integration, specifically focused on OpenAI and ElevenLabs, will be discussed by considering the views of 10 respondents. The main focus of this analysis is (1) the respondents' experience with OpenAI and Google Speech Recognition, and (2) their responses to the integration of OpenAI with ElevenLabs. In evaluating the use of the voice produced by Google Speech Recognition by being given audio samples of the results if OpenAI with Google's voice output, summarising their perceptions in 3 questions: (1) Opinions on the use of the voice produced by Google Speech Recognition get 70% of respondents said good, 10% very good, and 20% said less good. (2) The voice produced by Google Speech Recognition is felt to be more natural. 50% of respondents disagreed while the other 30% agreed, and the remaining 20% also said strongly agreed. Then, the last question (3) regarding the level of clarity of the voice produced by Google got a response of 70% of respondents expressing clear and 10% very clear. The remaining 20% said it was not clear. Table 1 presents respondents with answers to their perceptions of OpenAI with the results or voice output using Google Speech Recognition.

Table 1 Respondents' Results Regarding OpenAI & Google

| Respondent | Voice Usage Opinion | More Natural Sound | Voice Clarity Level |
|------------|---------------------|--------------------|---------------------|
| 1          | Less Good           | Disagree           | Unclear             |
| 2          | Very Good           | Strongly Agree     | Clear               |
| 3          | Good                | Agree              | Clear               |
| 4          | Good                | Agree              | Clear               |
| 5          | Good                | Strongly Agree     | Very Clear          |

|    |           |          |         |
|----|-----------|----------|---------|
| 6  | Good      | Disagree | Clear   |
| 7  | Good      | Agree    | Clear   |
| 8  | Good      | Disagree | Clear   |
| 9  | Good      | Disagree | Clear   |
| 10 | Less Good | Disagree | Unclear |

Furthermore, in evaluating the sound produced by ElevenLabs, which also involved 10 respondents who had been inserted with a sample of the audio output, 5 questions were given and a summary of the perceptions in each question: (1) Knowing or having used ElevenLabs before and the results were 80% of respondents had never used and 20% had knowledge of ElevenLabs. (2) Opinions about the sound quality produced by ElevenLabs and Google, 70% of respondents expressed that it was better than the sound quality produced by Google. 20% felt that the quality was comparable or the same and 10% felt that the quality was less good than Google's sound quality. The next question (3) asked whether the voice produced by ElevenLabs was human-like or not. 90% of respondents felt that ElevenLabs' voice was similar to a human voice and only 10% felt it was not similar enough. (4) Comfort with the use of voices from OpenAI and ElevenLabs in certain contexts such as virtual assistants. The results showed that 80% of respondents felt comfortable and the remaining 20% felt very comfortable. Then for the last question, (5) assessment of the clarity and readability of the voice from ElevenLabs with the results 90% of respondents said it was clear and 10% said it was very clear. Table 2 below shows the results of 10 respondents regarding the integration of OpenAI with ElevenLabs.

*Table 2 Respondents' Results Regarding OpenAI & ElevenLabs*

| Respondent | Ever used ElevenLabs? | Voice Quality (compared to Google) | Voice Similar to Human? | Use of voice in certain contexts, virtual assistants | Voice Clarity Level |
|------------|-----------------------|------------------------------------|-------------------------|------------------------------------------------------|---------------------|
| 1          | No                    | Better                             | Yes                     | comfortable                                          | Clear               |
| 2          | No                    | Same                               | Yes                     | comfortable                                          | Clear               |
| 3          | No                    | Better                             | Yes                     | comfortable                                          | Clear               |
| 4          | No                    | Less Good                          | No                      | comfortable                                          | Clear               |
| 5          | No                    | Better                             | Yes                     | comfortable                                          | Clear               |
| 6          | Yes                   | Better                             | Yes                     | comfortable                                          | Clear               |
| 7          | Yes                   | Same                               | Yes                     | comfortable                                          | Clear               |
| 8          | No                    | Better                             | Yes                     | comfortable                                          | Clear               |
| 9          | No                    | Better                             | Yes                     | Very comfortable                                     | Very Clear          |
| 10         | No                    | Better                             | Yes                     | Very comfortable                                     | Clear               |

From the evaluation results of respondents' perceptions of conversations with OpenAI with voice output by Google Speech Recognition and ElevenLabs, some interesting conclusions can be drawn. In Google, the majority of respondents gave a positive assessment of the quality of the voice produced. The level of clarity is also quite a positive point, although there are some respondents who consider the level of clarity to be low. Although most respondents felt that the sound was quite natural, there were still some who felt that there was room for improvement in the naturalness of the sound. As for ElevenLabs, the majority of respondents had no previous experience with ElevenLabs, suggesting that the AI technology may not be widely known. There were differing opinions on whether the voice quality was better, comparable or less than Google Speech Recognition but it was noted that many said it was better. ElevenLabs' voice was rated by most respondents as similar to the human voice and the clarity of the voice was clear.

Several respondents expressed their comfort level with the use of voice from both platforms, OpenAI and ElevenLabs, especially in the context of virtual assistants or text readers as comfortable and very comfortable, making concerns about safety and comfort of use likely to be important factors in further adoption. Thus, the results of this evaluation imply that the integration of OpenAI with ElevenLabs is better able to address the two main focuses in this research, namely identification in the context of audio or voice. (1) It overcomes the challenge of converting voice to text and vice versa. (2) The resulting voice will be human-like enabling a conversational experience similar to live communication with a real human. Although Google voice can also overcome the challenge of converting voice to text and vice versa, the voice output from Google cannot overcome the challenge of the resulting voice being human-like. Therefore, integrating OpenAI with ElevenLabs would be more suitable to develop a tool that has the potential to support learning with the concept of creating a conversation about learning with artificial intelligence technology that is similar to direct communication with real humans.

## Conclusion and Suggestion

This research achieved its goal by successfully integrating the two platforms OpenAI and ElevenLabs in the context of learning support. This successful integration translates into very realistic voice generation and will open up new potential in the development of tools that have the potential to support learning. The use of third-party platforms from OpenAI and ElevenLabs allows for realistic and effective voice interaction in education. In this discussion, we detail the advantages and limitations of each technology, including the realism of the voice produced by ElevenLabs and the natural language processing capabilities of OpenAI. The results show that ElevenLabs' technology is capable of producing voices that closely resemble real human voices, which can be used in various languages, including Bahasa Indonesia. However, some limitations, such as financial and technical limitations, need to be considered. The use of OpenAI to provide relevant natural language responses and its integration with ElevenLabs opens up opportunities to make voice-based learning interactive and informative. Using Bahasa Indonesia with the same voice can increase the personalization of interactions in an educational context. In further development, there is potential to add a "correct spelling" feature that can help correct English words or sentences. This would be a valuable addition to interactive learning with conversation, where artificial intelligence-based justification can help users improve their language skills. As technology develops, the combined use of OpenAI and ElevenLabs can be further enhanced to provide a more complete and effective educational solution.

## Reference

- Abraham, G., & Nithya, M. (2021). Multi-Functional Personal Assistant Robot Using Raspberry Pi and Coral Accelerator. *Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021*, 638–643. <https://doi.org/10.1109/ICCMC51019.2021.9418299>
- Akhtar, S., Shah, S. W. A., Rafiq, M., & Khan, A. (2016). Research design and statistical methods in Pakistan Journal of Medical Sciences (PJMS). *Pakistan Journal of Medical Sciences*, 32(1), 151. <https://doi.org/10.12669/PJMS.321.9033>
- Artificial intelligence (AI) | Definition, Examples, Types, Applications, Companies, & Facts | Britannica. (n.d.). Retrieved November 22, 2023, from <https://www.britannica.com/technology/artificial-intelligence>
- Asenahabi, B. M. (2019). Basics of Research Design: A Guide to selecting appropriate research design. *International Journal of Contemporary Applied Researches*, 6(5), 76–89. [www.ijcar.net](http://www.ijcar.net)
- Boru, T. (2018). CHAPTER FIVE RESEARCH DESIGN AND METHODOLOGY 5.1. Introduction. *CHAPTER FIVE RESEARCH DESIGN AND METHODOLOGY 5.1. Introduction*, December, 41. <https://doi.org/10.13140/RG.2.2.21467.62242>
- Carlini, N., & Wagner, D. (2018). Audio adversarial examples: Targeted attacks on speech-to-text. *Proceedings - 2018 IEEE Symposium on Security and Privacy Workshops, SPW 2018*, 1–7. <https://doi.org/10.1109/SPW.2018.00009>
- Character Education for the Digital Age. (n.d.). Retrieved November 22, 2023, from <https://www.ascd.org/el/articles/character-education-for-the-digital-age>

- Chen, H.-W. (2023). *Endoscopic Endonasal Skull Base Surgery for Pituitary Lesions: An AI-Assisted Creative Workflow to Develop an Animated Educational Resource for Patients and Physicians*. <http://jhir.library.jhu.edu/handle/1774.2/68519>
- Dipak Ghael, H., Solanki, L., Sahu, G., & Professor, A. (2020). A Review Paper on Raspberry Pi and its Applications. *International Journal of Advances in Engineering and Management (IJAEM)*, 2, 225. [https://www.researchgate.net/publication/348993428\\_A\\_Review\\_Paper\\_on\\_Raspberry\\_Pi\\_and\\_its\\_Applications](https://www.researchgate.net/publication/348993428_A_Review_Paper_on_Raspberry_Pi_and_its_Applications)
- Durga Prasad, K. Y., Alekhya, S., Naresh, A., & Rajesh, K. V. N. (2018). Voice Recognition Based Home Automation using Raspberry Pi. In *International Journal of Innovative Science and Research Technology* (Vol. 3, Issue 7). [www.ijisrt.com](http://www.ijisrt.com)
- ElevenLabs: *AI Voice Generator & Text to Speech*. (n.d.). Retrieved November 22, 2023, from <https://elevenlabs.io/about>
- Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Medical Education*, 9. <https://doi.org/10.2196/45312>
- Grimmett, R. (n.d.). *Raspberry Pi robotic projects: work through a mix of amazing robotic projects using the Raspberry Pi Zero or the Raspberry Pi 3*.
- Hill, C. (2016). Learning Scientific Programming with Python. *Learning Scientific Programming with Python*. <https://doi.org/10.1017/CBO9781139871754>
- Raje, S., Lagu, S., Joshi, G., Dolas, A., Lonkar, P. S., Raje, S., Lagu, S., Joshi, G., Dolas, A., & Lonkar, P. S. (2021). Speech Recognition using AI. *JETIR*, 8(7), b433–b439. <https://www.jetir.org/view?paper=JETIR2107182>
- Schantz, D., & Lindeman, C. A. (2019). The research design. *Journal of Nursing Administration*, 12(2), 35–41. <https://doi.org/10.1097/00005110-198202000-00007>
- Tuma, F. (2021). The use of educational technology for interactive teaching in lectures. *Annals of Medicine and Surgery*, 62, 231–235. <https://doi.org/10.1016/J.AMSU.2021.01.051>
- Vashistha, P., Singh, J. P., Jain, P., & Kumar, J. (2019). Raspberry Pi based voice-operated personal assistant (Neobot). *Proceedings of the 3rd International Conference on Electronics and Communication and Aerospace Technology, ICECA 2019*, 974–978. <https://doi.org/10.1109/ICECA.2019.8821892>