Journal of Applied Research and Technology

www.jart.icat.unam.mx

Journal of Applied Research and Technology 23 (2025) 155-163

Original

Implementation of an Indonesian AI-based text-to-speech system for self-student pickup announcements based on natural language processing

A. P. Atmaja* • S. K. Bimonugroho • M. H. R. Ismar

State Polytechnic of Madiun, East Java, Indonesia

Received 06 07 2024; accepted 10 07 2024
Available 04 30 2025

Abstract: Artificial intelligence (AI) and natural language processing (NLP) technologies are advancing rapidly, offering substantial assistance in various fields. This study focuses on developing a text-to-speech (TTS) device tailored for elementary schools in Indonesia to autonomously call students during pickup times. Utilizing the Google text-to-speech (gTTS) library and Python, the device operates on a low-specification Mini PC. It integrates with a cloud-based student management information system (MIS) to synchronize student data and announcements. The device automates the student call-out process, reducing the workload of school staff and ensuring clear, accurate pronunciation. Successfully tested, the device demonstrates a practical, cost-effective solution for modernizing student pickup systems. It showcases the potential of AI and NLP in educational environments, emphasizing operational efficiency and technological accessibility.

Keywords: Self-student caller, autonomous pickup announcements, AI, NLP, text-to-speech, gTTS.

*Corresponding author.
E-mail address: atmaja@pnm.ac.id (A. P. Atmaja).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

1. Introduction

The world is constantly changing towards the era of automation and personal assistance. Artificial intelligence (AI) technology is increasingly helping humans to do something. One branch of artificial intelligence is natural language processing (NLP). NLP is a set of methods for making human language accessible to computers (Eisenstein, 2018). AI and NLP are being used largely in the world. It enables communication between people and computers and automatic translation to facilitate easy interaction with others around the world (Natural Language Processing in Artificial Intelligence, 2020). NLP techniques as an AI approach have been leveraged to extract information from clinical narratives in electronic health records (Juhn & Liu, 2020). It can also be applied to address many of the information needs made urgent by the COVID-19 pandemic. This review surveys approximately 150 NLP studies and more than 50 systems and datasets addressing the COVID-19 pandemic (Chen et al., 2021). Another study aims to examine how the Internet of Things (IoTs) connects text objects as part of NLP and AI responding to human needs (Mah et al., 2022). In NLP technology, computers are given the ability to understand natural language from humans, such as writing in the form of letters or spoken sounds. The use of NLP technology can be found on many devices, such as translators, sentiment analysis, chatbots, text processing, digital assistants, text-to-speech conversion software, etc. A useful text-to-speech synthesizer in the form of a simple application has been developed. It converts inputted text into synthesized speech and reads out to the user which can then be saved as an mp3 file (Mukherjee et al., 2018).

Text-to-speech (TTS) is a technology that allows computers to convert text into sounds or speech that can be heard by humans. In TTS, the computer reads and converts text into speech based on certain grammatical and syntactic rules. With TTS technology, computers can produce sounds that are almost the same as human voices and help users access text in the form of voice. TTS is often used in screen reader applications, digital assistants, e-book reading devices, remote calling devices, and other devices that require voice interaction with the user. The broadcasting service application by implementing text-to-speech on a web-based radio broadcasting service application has been developed. It can convert data/information in text form into information in the form of voice (Sudiartha et al., 2022). Another paper is to shed light on the impact of the shadowing technique with the aid of spoken text features by Google text-to-speech system on removing learners' flat tones as well as achieving basic English intonation in the text (Lee et al., 2022). While some study combines the concept of Optical Character Recognition (OCR) and Text to Speech Synthesizer (TTS) in Raspberry Pi. This kind of system helps visually impaired people to interact with computers effectively through a vocal interface (Lakshmi et al., 2016).

In elementary schools, it has become common to have loudspeakers used to give announcements to students. In the case of picking up students' home, the call of students is carried out conventionally, which is done repeatedly by the officer when the pick-up is waiting. This requires an officer or operator and a set of microphones, so it becomes a very tiring job for operators to call students who are playing in the inner courtyard of the school to come out to their respective pick-ups. Therefore, there is a need for a solution to be able to develop an independent student summoning device to become an assistant and even replace the role of the calling officer, as well as so that the pronunciation of the name of the student being called becomes easier to understand and attracts students to listen to it. This research introduces a novel application of TTS technology specifically tailored for school environments, which has not been widely explored in previous studies. The integration of gTTS with a portable Mini PC platform for this specific purpose presents an innovative approach to modernizing the traditional student pickup system. The developed device aims to significantly reduce the workload of school staff by automating the student call-out process, thus enhancing efficiency and accuracy. Additionally, by utilizing low-specification hardware, this research demonstrates a cost-effective solution that can be easily adopted by schools with limited technological resources. This work not only showcases the practical implementation of NLP and TTS in a real-world setting but also sets a precedent for further applications of AI in educational environments.

2. Methods

The stages in the method used include study literature, analysis, design, and implementation.

2.1. Study literature

Several studies have conducted somewhat similar experiments. The first research focuses on natural language processing (NLP). NLP is a branch of computer science that focuses on the computational processing of human language or natural language. NLP involves using technologies and methods such as speech recognition, voice synthesis, and machine translation to help computers understand, process, and generate human language automatically and more effectively (Jurafsky & Martin, 2008). This is a processing method that helps computers to understand the natural language of human beings and hence use it in a way to optimize the area of artificial intelligence (Dubey, 2022). It has spread its applications in various fields such as machine translation, email spam detection, information extraction,

summarization, medical, and question answering, etc. (Khurana et al., 2023). NLP involves several stages, such as tokenization (breaking down text into the smallest units such as words or phrases), recognition and analysis of grammatical structures, as well as extraction of information from text. In addition, NLP also involves the use of language models, machine learning, and big data processing technologies to improve system performance. Figure 1 shows the main elements of NLP consisting of computer science, human language, and artificial intelligence.

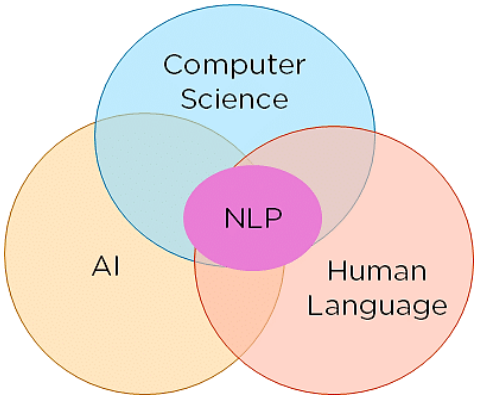

Figure 1. Main elements of NLP.

A study has investigated the impact of text-to-speech (TTS) on reading comprehension in children with reading difficulties. The results were that reading comprehension was significantly higher for the TTS condition compared to the No TTS condition. Thus, TTS is a useful compensatory reading aid to improve comprehension (Keelor et al., 2020). In education, applications with text-to-speech (TTS) and Speech-to-Text (STT) functions have been widely used as Assistive Technology (AT) for students in schools with reading difficulties. A portable AT developed in one study has the potential to improve reading skills in regular education participation (Nordström et al., 2019). Sulaimon et al. have also developed a TTS device to help the reading comprehension scores of two fourth-grade students with learning disabilities in an urban school. The results showed that the TTS program was easy to use with little or no support (Sulaimon & Schaefer, 2023). Another study examined the impact of a text-to-speech program on the academic achievement of high school students with disabilities in an online transition curriculum that emphasized information literacy. Findings showed that text-to-speech support improved unit quizzes and reading comprehension performance with large effect sizes (Izzo et al., 2009). Even an online application for the blind that combines speech-to-text, non-visual desktop access, and text-to-speech is one of the drivers of the fulfillment of inclusive education needs and services on campus based on the educational needs and special services for blind students (Perianto et al., 2021). The development of text-to-speech has also been applied in administrative matters. A framework was developed to automate the tedious task of writing thousands of emails manually, where the voice interaction of the sender's name, subject, and email body is made into an audio-controlled email bot (Shah et al., 2021).

NLP technology is often used in Internet of Things (IoT) devices. There was a study focused on the design and implementation of a speech recognition system integrated with IoT to control electrical appliances and doors with Raspberry Pi as a core element (Abdulkareem et al., 2021). The system composed of IoT and cloud systems is also used as a real-time monitoring system for aquaculture and is supported by NLP-based chatbot assistants to facilitate the user (Al Rasyid et al., 2020). NLP and IoT are also can be integrated to enhance security. A system for identifying malicious behaviors in an IoT environment using n-gram and word-embedding techniques has been developed (Kumar & Kumar, 2023). While a bus arrival information system using IoT devices and Google text-to-speech system with Firebase Cloud Messaging has also been developed (Aryotejo et al., 2020). Another study proposed a speech recorder and translator to solve the problem of translation between different languages using Google Cloud Speech-to-Text (Wang, 2021).

2.2. Analysis and design

At the analysis stage, an analysis of the initial problems found in the problem domain will be carried out, namely, how to make a device that can produce sounds to call a student in elementary school based on a database of student names that have existed before on the cloud server. The device utilizes NLP technology in the form of converting writing into voice or text-to-speech. The device will later be installed on the school grounds that can be accessed by parents or pick-up their children easily. We used Google text-to-speech system as a cloud-based API that converts text into natural-sounding speech. It uses a variety of artificial intelligence (AI) techniques, including phonetics, pronunciation dictionaries, and voice models.

Phonetics is the study of the sounds of human speech. Google text-to-speech system uses phonetics to break down the text into its individual sounds. Pronunciation dictionaries contain the pronunciation of words in different languages. Google text-to-speech system uses these dictionaries to determine how to pronounce each word. While voice models are trained on large amounts of audio data to learn how to produce natural-sounding speech, they are used to generate the audio output. In addition to these techniques, Google text-to-speech system also uses other factors to improve the quality of the speech output, such as speech rate, pitch, and volume. Speech rate allows us to control the speech rate to

suit our needs. Pitch is the highness or lowness of speech. It is used to control the pitch to create a variety of effects. And volume is to control the volume to make the speech easier to hear. The text processing process in Google text-to-speech system involves the steps shown in Figure 2.

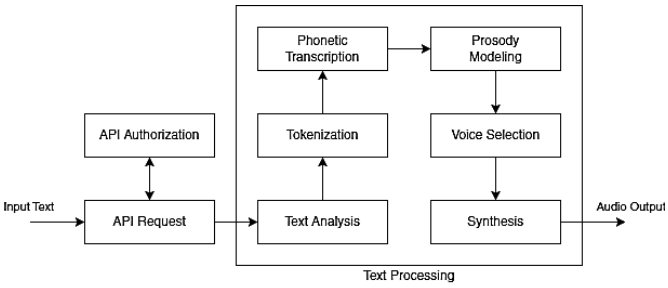

Figure 2. Google text-to-speech system's flow.

The process begins with a user or application providing a text input. This text in the Bahasa Indonesia language was taken from the student's database. A request is formulated to the Google text-to-speech API, including the text that wants to convert to speech, desired voice settings, and other parameters, and the API credentials to authorize the request. This ensures that only authorized users or applications can access the service. The input text is analyzed to understand its structure, grammar, and context. This analysis helps determine how individual words and sentences should be pronounced, as well as where to add pauses and intonation. The text is first tokenized, which means it is divided into individual words or syllables. For each token, the TTS engine identifies the phonetic transcription. This is a sequence of phonemes, which are the smallest units of sound that can distinguish one word from another in a language. The phonetic transcription of a word is typically determined using a dictionary or a statistical model. On the Prosody modeling, the TTS engine then models the prosody of the speech, which includes the pitch, loudness, and rhythm of the speech. The TTS engine then selects a voice that matches the desired gender, accent, and other characteristics and then synthesizes the speech sounds. The synthesis process is the final step in the text-processing process. In this step, the TTS engine generates the speech sounds based on the phonetic transcription, prosody model, and selected voice. The synthesis process is typically carried out using a vocoder, which is a device that generates sound from a digital signal.

At the design stage, an application design will be made which includes data structures, software architecture, detailed procedures, and interface characteristics. This design process aims to make the software understandable before starting to write the program. Figure 3 shows the architecture of the system proposed. The student caller device will connect to the internet to call the student database managed by the student management information system (MIS) stored on the State Polytechnic of Madiun (PNM) cloud server. This Student MIS will be fully managed by the school to manage the data of student names that can be called by name by the student caller device. A Google text-to-speech library will be used as an NLP algorithm on this device. Student data on the cloud server will be synchronized with the database on the student's caller device that has been installed. It includes synchronizing the names of students and also recording the history of the day and time the students were called. So that everything will be fully documented. It is expected to be more intuitive in pronunciation in calling the names of students in Indonesian.

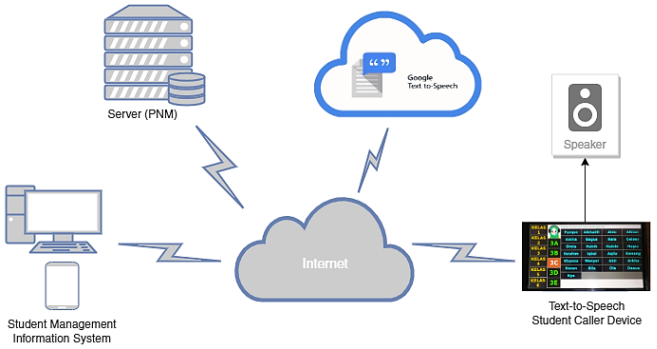

Figure 3. System architecture.

The text-to-speech based student caller system has a flowchart consisting of four entities, namely Admin, School Operator, Student Pickup, and System. The system includes IoT hardware and software connected to a cloud server. Figure 4 shows the flowchart of the system proposed. Admin is the owner of this system platform which only consists of one user. School operators are schools that use this device. Meanwhile, student pickups are people who use this device to call students from grades 1 to 6 to be picked up home. Meanwhile, the system is a series of devices themselves, both software and supporting hardware.

Description of the system flowchart stages as follows:

1) The initial stage that must be carried out by Admin, is to add school data as a device user. This must be done so that this system can be used by schools. This school account will be used by the school to manage student data that can be called using this device.
2) After the school gets its account, the school (operator) can manage students via a browser using the URL address https://panggil.research.pnm.ac.id. Management can include adding, changing, deleting, and searching for students. At this stage, the data managed by the school operator will be stored in a cloud database.
3) The Bot program stored on the calling device that is installed at the school will periodically synchronize

the cloud database and ensure that the data stored in the cloud database will be the same as the database stored in the device database. Both databases use MySQL. Hence, the information displayed on the student caller device will be the same as the information on the student management website.

Student pick-ups can call the students they want to pick up by selecting a nickname via the touch screen on the device themself. Once a name is selected, the loudspeaker will sound the name. This happens because the artificial intelligence (AI) program on the device being developed has translated the text of the name stored in the database into a sound with the same intuition as that spoken by humans.

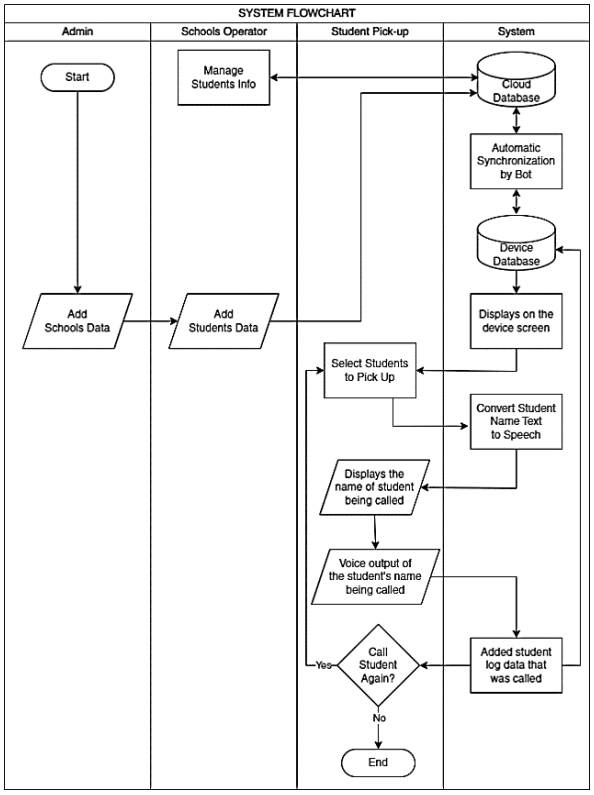

Figure 4. System's flowchart.

2.3. Implementation

To implement this device requires some hardware. After the hardware is assembled, a program is then given to provide the desired instructions. This developed device implements NLP-based TTS technology to call students via loudspeaker who have been picked up home by the pick-up students independently at school using a touchscreen by utilizing the Google text-to-speech (gTTS) library using the Python 3.8 programming language and runs on a Mini PC platform with Intel Atom Z3735F 1.33Ghz CPU specifications, 2GB DDR3 RAM, 32GB eMMC Storage, HDMI Video & Audio Output, 3.5mm Audio Output, Networking AP6212 Support IEEE 802.11b/g/n, Fast Ethernet 10 /100Mbps. The results of application design on touchscreen devices are shown in Figure 5.

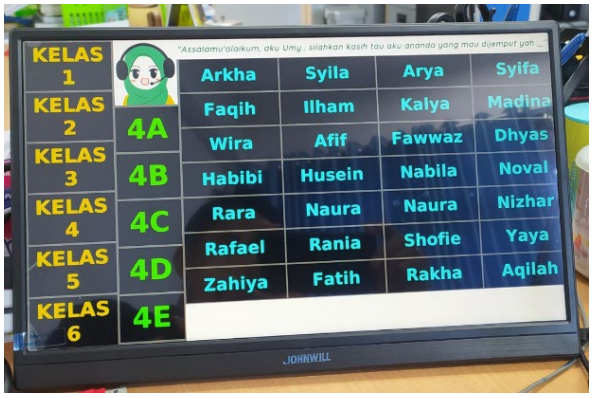

Figure 5. Application design results on touchscreen devices.

Web-based applications are used for the administration and management of student data that will be used by student caller devices. All web-based applications are built using the Laravel version 9 framework using the PHP programming language version 8.0. The interface of this application is made in such a way that it can be accessed well via PC, tablet, or smartphone. Only certain users can access this web application. The server for this system is located at PNM which can be accessed using the URL https://panggil.research.pnm.ac.id.

3. Results and discussion

The self-student caller device has been successfully installed at the school and has been tested. This device is connected to the student management system on the server, so the data is synchronized. The synchronization process occurs periodically between the server and the device, according to the flowchart in Figure 3. This device can be installed on the wall in a place that is easily accessible to those who pick up their children when they come home from school so that those who pick them up can access this device independently to call the students they want to pick up. Figure 6 shows the design for placing the device on a wall which will later be protected with iron brackets to maintain the safety of the device when not in use. The bracket has two security locks, namely a lock on the front side of the screen for daily use of the device and a lock on the back side to lock the back of the device which consists of the electrical circuit and the Mini PC. Figure 7 shows a picture of the back of the device.

Figure 6. Result of device's implementation.

Figure 7. Back side of the device.

On the student's management website, which can be accessed at the URL https://panggil.research.pnm.ac.id, it can only be accessed by previously registered users. Below are shown pictures of the display of student management on the student manager's website. Figure 8 is a history of student calls that have been made. Users can see detailed information per student. While Figure 9 is a chart of student calls by day.

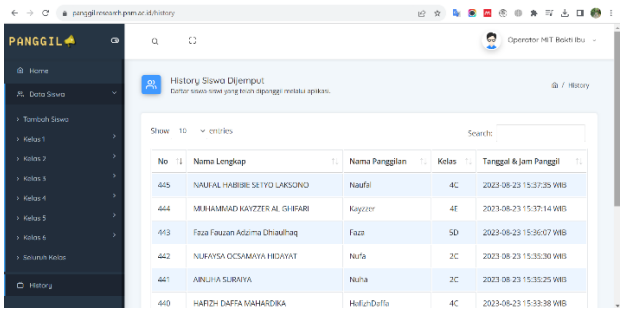

Figure 8. Student history called webpage (school operator page).

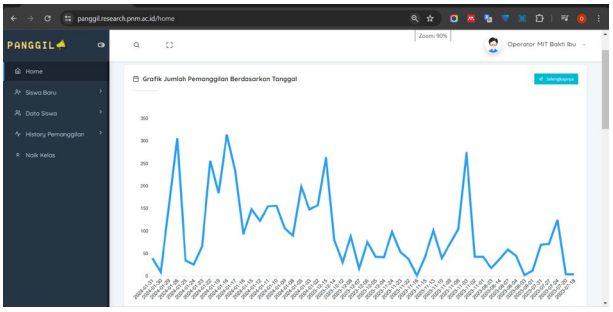

Figure 9. Number of daily call charts (school operator page).

Text-to-speech-based self-caller natural language processing (NLP) devices have been used to call students independently by schools that use them (Atmaja et al., 2023). This device has communicated well with the server located at PNM. The data stored on the PNM server is the same as the data stored on the device. The system synchronizes student data and student call history on school devices and servers periodically, namely every 10 seconds. This is done by a Bot that has been planted on the device. Therefore, the synchronization process only occurs when the device installed at the school is on. The system can change the text of student's names that have been stored in the database into sound properly so that the process of calling students using connected loudspeakers can be heard clearly. Figure 10 below shows the devices used by the school.

Figure 10. Use of the device (touchscreen).

This implementation of a TTS-based caller device presents several advantages and unique features. Similar research by Keelor et al. (2023) on TTS technology in educational environments focused on reading assistance rather than administrative tasks. While their system demonstrated the utility of TTS for learning enhancement, it did not address the logistical aspect of student management as our device does. Another study by Meurers (2020) utilized NLP for interactive educational tools, primarily aimed at language learning. That study underscored the importance of real-time text-to-speech

conversion but did not explore the potential of such technology for operational efficiency in schools. Our device differentiates itself by integrating seamlessly with the existing student management system and providing real-time updates and synchronization. The synchronization interval of every 10 seconds ensures that the data remains current and accurate, which is critical for operational reliability. Additionally, the device's implementation on a low-specification Mini PC platform makes it a cost-effective solution compared to the higher-end hardware required in other studies. Moreover, the use of the Google text-to-speech (gTTS) library within a Python programming environment highlights the accessibility and adaptability of open-source tools in creating practical AI applications. This approach not only reduces development costs but also allows for easier modifications and updates, as compared to proprietary solutions used in previous research.

4. Conclusions

A self-student caller device that implements text-to-speech based on natural language processing (NLP) has been successfully developed. This system is completely dependent on the internet because it uses a server that runs online and is dependent on the online libraries used. Users can independently choose the name of the student they will pick up via a 15.6-inch touchscreen that has been installed with an NLP device based on Google text-to-speech system using the Python programming language that has been developed. As a result, the name text in the database can be converted into sound by the application well according to the Indonesian language in general. This device will continue to communicate with the server asynchronously for a certain period time. In managing it, users will access the student management information system website online from anywhere and at any time, and the Bot program that has been developed will synchronize when the devices installed in the school are turned on. The self-student caller device not only provides a novel application of TTS technology in the administrative context but also sets a benchmark for future developments in the use of AI in school environments. By offering a practical, cost-effective solution that enhances operational efficiency, this research contributes significantly to the field and opens up new avenues for further innovation in educational technology.

The device produced from this research has an impact on increasing the competitiveness of the school, where parents will be more satisfied in sending their children to schools that provide this independent student calling service. The implementation of this device can also create a more orderly school atmosphere, especially during homecoming hours. This provides convenience for parents because it reduces the potential for confusion or errors when picking up children, which can also increase student safety in the school environment. Meanwhile, the impact on the national education strategy, this TTS device contributes to the government's vision to integrate technology into the national education system. The use of this technology also supports the Indonesian government's program to accelerate digital transformation in various fields, including education. With its low cost and minimal technical specifications, this tool makes it easier for schools with limited budgets to adopt new technology.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgments

We would like to thank the Department of Information Technology and Department of Software Engineering Technology for their support and the laboratory to conduct this research, as well as the institution for the opportunity given to us.

Funding

The authors have received funding support for this research from the State Polytechnic of Madiun, Indonesia, through a higher competitive grant scheme with the number of DIPA 023.18.2.677632/2023.

References

Abdulkareem, A., Somefun, T. E., Chinedum, O. K., & Agbetuyi, F. (2021). Design and implementation of speech recognition system integrated with internet of things. *International Journal of Electrical and Computer Engineering*, *11*(2). https://doi.org/10.11591/ijece.v11i2.pp1796-1803

Al Rasyid, M. U. H., Sukaridhoto, S., Dzulqornain, M. I., & Rifa'i, A. (2020). Integration of IoT and chatbot for aquaculture with natural language processing. Telkomnika (*Telecommunication Computing Electronics and Control*), 18(2). https://doi.org/10.12928/TELKOMNIKA.V18I2.14788

Aryotejo, G., Hidayatulloh, A. S., & Mufadhol, M. (2020). The Bus Arrival Information System for Tunanetra Based on Firebase Cloud Messaging and Google Text to Speech. *Journal of Applied Informatics and Computing*, *4(*2). https://doi.org/10.30871/jaic.v4i2.2102

Atmaja, A. P., Bimonugroho, S. K., Septianto, T., Aji, M. B., Al. Aziz R, M. S., & Saputra, D. A. (2023). Implementasi perangkat text-to-speech pemanggil siswa mandiri berbasis natural language processing di mit bakti ibu kota madiun. *Journal Hilirisasi IPTEKS*, *6*(4). https://doi.org/10.25077/jhi.v6i4.704

Chen, Q., Leaman, R., Allot, A., Luo, L., Wei, C. H., Yan, S., & Lu, Z. (2021). Artificial Intelligence in Action: Addressing the COVID-19 Pandemic with Natural Language Processing. In *Annual Review of Biomedical Data Science* (Vol. 4). https://doi.org/10.1146/annurev-biodatasci-021821-061045

Dubey, A. (2022). A Review Study on Computational Linguistics and Natural Language Processing. International Journal for Research Publication and Seminar, 13(2), 106–113. Retrieved from https://jrps.shodhsagar.com/index.php/j/article/view/578

Eisenstein, J. (2018). Natural language processing. *Jacob Eisenstein*, *507*.

Izzo, M. V., Yurick, A., & McArrell, B. (2009). Supported eText: Effects of Text-to-Speech on Access and Achievement for High School Students with Disabilities. *Journal of Special Education Technology*, *24*(3), 9–20. https://doi.org/10.1177/016264340902400302

Juhn, Y., & Liu, H. (2020). Artificial intelligence approaches using natural language processing to advance EHR-based clinical research. *Journal of Allergy and Clinical Immunology*, *145*(2), 463–469. https://doi.org/10.1016/j.jaci.2019.12.897

Jurafsky, D., & Martin, J. H. (2008). Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition (Vol. 23). Upper Saddle River: Pearson/Prentice Hall.

Keelor, J. L., Creaghead, N. A., Silbert, N. H., Breit, A. D., & Horowitz-Kraus, T. (2023). Impact of text-to-speech features on the reading comprehension of children with reading and language difficulties. Annals of dyslexia, 73(3), 469-486. https://doi.org/10.1007/s11881-023-00281-9

Keelor, J. L., Creaghead, N., Silbert, N., & Horowitz-Kraus, T. (2020). Text-to-speech technology: Enhancing reading comprehension for students with reading difficulty. *Assistive Technology Outcomes & Benefits*, *14*(1), 19-35.

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications*, *82*(3). https://doi.org/10.1007/s11042-022-13428-4

Kumar, Y., & Kumar, V. (2023). Security in IoT systems using natural language processing: Future challenges and directions. *In Internet Technology Letters* (Vol. 6, Issue 4). https://doi.org/10.1002/itl2.411

Lakshmi, K., Lakshmi, K., & Rao, Mr. T. C. S. (2016). Design and Implementation of Text To Speech Conversion Using Raspberry Pi. *Ijitr*, *4*(6).

Le, P. N., Vu, H. M. L., & Tran, M. N. (2022). Improving EFL students' intonation in-text using shadowing technique with the implementation of Google text-to-speech. *AsiaCALL Online Journal*, *13*(1), 93-121. https://asiacall.info/acoj/index.php/journal/article/view/102

Mah, P. M., Skalna, I., & Muzam, J. (2022). Natural Language Processing and Artificial Intelligence for Enterprise Management in the Era of Industry 4.0. *Applied Sciences (Switzerland)*, *12*(18). https://doi.org/10.3390/app12189207

Meurers, W. D. (2020). Natural language processing and language learning. https://doi.org/10.1002/9781405198431.wbeal0858.pub2

Mukherjee, P., Santra, S., Bhowmick, S., Paul, A., Chatterjee, P., & Deyasi, A. (2018). Development of GUI for text-to-speech recognition using natural language processing. *2018 2nd International Conference on Electronics, Materials Engineering and Nano-Technology, IEMENTech 2018*. https://doi.org/10.1109/IEMENTECH.2018.8465238

Natural Language Processing in Artificial Intelligence. (2020). In Natural Language Processing in Artificial Intelligence. *Apple Academic Press*. https://doi.org/10.1201/9780367808495

Nordström, T., Nilsson, S., Gustafson, S., & Svensson, I. (2019). Assistive technology applications for students with reading difficulties: special education teachers' experiences and perceptions. Disability and Rehabilitation: *Assistive Technology*, *14*(8), 798–808. https://doi.org/10.1080/17483107.2018.1499142

Perianto, E., Rianto, R., Pranowo, T. A., Noormiyanto, F., Hidayat, L., & Ciptadi, P. W. (2021). Pengembangan Aplikasi Pembelajaran Tunanetra (APTUN) Berbasis Teknologi Asistif untuk Pencarian Konten Pembelajaran Mahasiswa Tunanetra. Elementary School: *Journal Pendidikan dan Pembelajaran Ke-SD-An*, *8*(1). https://doi.org/10.31316/esjurnal.v8i1.1208

Shah, H. D., Sundas, A., & Sharma, S. (2021). Controlling Email System Using Audio with Speech Recognition and Text to Speech. *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions*), ICRITO 2021. https://doi.org/10.1109/ICRITO51393.2021.9596293

Sudiartha, I. K. G., Prihatini, P. M., & Purbhawa, I. M. (2022). Implementation of text to speech on web-based broadcasting radio service applications. Matrix: *Journal Manajemen Teknologi Dan Informatika*, *12*(1), 51–57. https://doi.org/10.31940/MATRIX.V12I1.51-57

Sulaimon, T., & Schaefer, J. (2023). The Impact of Text-to-Speech on Reading Comprehension of Students with Learning Disabilities in an Urban School. *TechTrends*, *67*(2), 376–383. https://doi.org/10.1007/s11528-022-00800-2

Wang, H. H. (2021). Speech Recorder and Translator using Google Cloud Speech-to-Text and Translation. *Journal of IT in Asia*, *9*(1). https://doi.org/10.33736/jita.2815.2021