Implementation of Google Translate Application Programming Interface (API) as a Text and Audio Translator

3 authors:

Zhara Nabila
Lampung University
2 PUBLICATIONS   **7** CITATIONS

SEE PROFILE

Humairoh Ratu Ayu
Lampung University
13 PUBLICATIONS   **95** CITATIONS

SEE PROFILE

Arif Surtono
Lampung University
54 PUBLICATIONS   **167** CITATIONS

SEE PROFILE

Implementation of Google Translate *Application Programming Interface* (API) as a Text and Audio Translator

Zhara Nabila¹*, Humairoh Ratu Ayu², Arif Surtono³
¹²³Physics, Faculty of Mathematics and Natural Sciences,
University of Lampung, Indonesia

Abstract .*Through this translator program, it is craved that it can avail the general public to understand foreign language videos, can be useful in the world of education and technology, and can avail the persons with disabilities be up to communicate.* The method used is a classification method that functions to detect the flow of shapes to instruct the class attribute as the task of the input attribute by generating automatic output through three stages, namely Machine Learning, Natural Language Processing, and Speech. *The results showed that 90.38% of videos were successfully translated into text and audio, 9.62% of videos failed to be translated because the owner limited public interaction, and 89%-97% synchronization between text and audio.* In this research, a text and audio translator program has been created using the Application Programming Interface (API). This program is a configuration of deep learning, machine translation, and text-to-speech designed using the high-level programming language python. The system used is a predictive system in which the system tries to predict the output equally the wishes of the user.

Keywords: *Application Programming Interface, Artificial Intelligence, Persons with Disabilities, Python Language, Information Technology*

INTRODUCTION

Every person needs a medium in interacting. The medium is language. Language is a means of individuals interacting in the form of spoken sound symbols by the medium of human speech [1]. In understanding language, an object study is needed, namely linguistics. The origin of the word linguistics is adopted from the *lingua* which means language. Linguistics is the diagnosis of human language with science [2].

The character of a nation can be seen from the use of language. As a multilingual nation, Indonesia has hundreds of regional languages. To facilitate people interaction, one language is set as a unifier of the nation. This event has been arranged in the state constitution, the 1945 Constitution in Article 36, in which Indonesian is the official language of the state [3].

Recent research shows, that worldwide there are 7,000 languages and according to Ethnology statistics, 2.7 billion people communicate with eight languages and about 70 million humans interact with 5,727 languages [4]. With so many foreign languages as national languages, various problems will arise, either in the form of economization or efficiency. As is the case with the Philippines, although the Philippines has a national anthem (Lupang Himirang) in terms of language, the Philippines is unable to establish Tagalog as a unifying language [5].

With the globalization era, there have been changes in the times, including the linguistic aspects in it. To answer this question, English is designated as a global language with more than one billion people speaking it. The global language can broaden horizons in the field of science and technology. The dissemination of up-to-date information can help every individual around the world in getting information [6].

English has been established as a universal language with a limited understanding of a language meaning, it becomes a big enough problem for someone to communicate. A simple example is when watching YouTube videos that use foreign languages, of course, it is very difficult to understand these videos,

especially since not all YouTube videos provide subtitles and dubbing in them. Therefore, a translator program is implemented by utilizing the Application Programming Interface (API) which is a management in the form of interfaces, tasks, levels, to forms to design software or computer software. Application Programming Interface (API) can help programmers "manifest" software that is integrated with other devices. Application Programming Interface (API) is a connector between applications that allows programmers to use disposition functions that are managed through a computer program operation [7].

This research refers to the lack of development in linguistics through Artificial Intelligence by utilizing computer programs. Then the Google Translate Application Programming Interface (API) was implemented as a text and audio translator with the main goal of making this translator program, which is expected to help the general public understand foreign language videos so that they add insight, can be useful in the world of education and technology is knowing how Artificial Intelligence works. Intelligence and Programming, and can help people with disabilities to be able to communicate audio and visually.

METHODS

The system used is a predictive system where the system tries to predict the output in accordance with the wishes of the user by utilizing the available libraries. An Application Programming Interface (API) is the design of a task that other programs can invite. The Application Programming Interface (API) functions as a link that unifies various applications from each version of the manifesto [8]. The Application Programming Interface (API) can translate foreign languages, such as English to Indonesian into script text and can project appropriate audio translations.

This program uses several libraries, namely Google API, Google Translate, and Google TTS. TTS or text-to-speech is an implementation with the aim of converting from text to audio form where the resulting conversion can be adjusted for prosody speed [9]. The Application Programming Interface (API) forwards an invitation to the system in the form of translation evidence and the system displays the appropriate translation response [10].

The method used is a classification method in which each state stores various attributes, such as class attributes. This method works by detecting a stream of shapes to instruct class attributes as an assignment from the input attribute by generating the appropriate output via a language video link from YouTube. To find some class attributes, you can use a simpler programming language. Through python, the input will be more concise with the output remaining the same. Python programming language is a language that accommodates level automatic programming that can be run in various programs by managing programs quickly with simple encryption, supporting various platforms, and having a memory management disposition [11]. To understand the creation of an Application Programming Interface (API) translator program, a program creation scheme is shown.
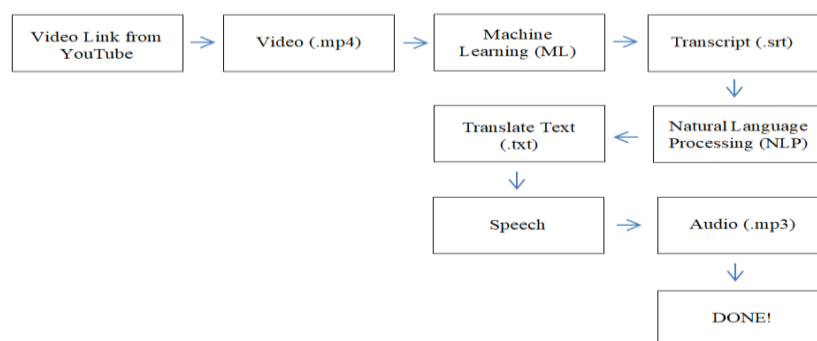
Figure 1. Flow Chart of Application Programming Interface (API) Programme

On the Jupiter notebook working page in the python application, a video link is inserted into the program, which is connected to the internet network because the created program must be connected to a stable internet connection. Make sure beforehand that you have downloaded the libraries that will be used. Libraries used are Google API, Google Translator, and Google Text-to-Speech.

The first stage is a deep learning process that is included in the Machine Learning (ML) component. Deep Learning is a subset of Artificial Intelligence (AI) that focuses on big neural network modeling, which is capable of making accurate data-driven decisions. In digital era, many high-end consumers and online technology companies use deep learning, such as Facebook which uses it to processing text in online conversations and Google, Baidu, Microsoft, etc. use it for search image and translator [12]. The system will convert the video into a script by utilizing the libraries from the Application Programming Interface (API).

The second stage is the machine translation process which is one of the constituents of Natural Language Processing (NLP). Machines will interact by utilizing natural language. Natural language is the language used in general communication that includes all mother tongues from all over the world. Natural Language Processing (NLP) is a very onerous duty in computer science where it presents an extensive variety of problems that vary from language to language. Foregoing, computer scientists broke down language into grammatical forms to become parts of speech, phrases, and others using complex algorithms [13]. Natural Language Processing (NLP) contains many algorithms, tasks, and problems in the form of input text and produces an output of useful information, such as labels, semantic representations, and so on. Other tasks in the form of translating, summarizing, and making texts are included in the form of Natural Language Processing (NLP) [14]. After the video script is obtained in the first stage, it needs to be converted into the desired translation script by making use of the Google Translate library. Google Translate is useful to answer users' needs for foreign language translations from various languages in the world.

The results of the translated script need to be converted into audio form first by using the Google Text-to-Speech library. This stage is the last stage of the Application Programming Interface (API) translator program, namely Speech. The machine will work according to the commands entered in the form of text-to-speech or speech-to-text and the output from the system will not only appear on the monitor but will be stored automatically on the user's device.

RESULT AND DISCUSSION
The program is tested by translating various languages on the video with output in the form of text and audio with a duration of 3 minutes to 12 hours.

Table 1. Results of Language Test and YouTube Video Duration

Link Video	1st Test	2nd Test	1st Result		2nd Result	
	Language	Duration	Text	Audio	Text	Audio
gdZLi9oWNZg	Korea	3:43	✓	✓	✓	✓
IHENlg8Se7M	Jepang	4:03	✓	✓	✓	✓
Fwf45pIAtM	Arab Maroko	4:14	✓	✓	✓	✓
71hi9H6fZuc	Arab	4:17	✓	✓	✓	✓
kLGCtPa4BEQ	Mandarin	4:36	✓	✓	✓	✓
kJQP7kiw5Fk	Spanyol	4:41	✗	✗	✗	✗
m8MfJg68oCs	Korea	4:42	✓	✓	✓	✓
bZU7qSZG1T0	Italia	5:06	✗	✗	✗	✗
clV1IzcSVQ4	Prancis	5:18	✓	✓	✓	✓
7C2z4GqqS5E	Korea	5:19	✓	✓	✓	✓
x44bYx-vPS4	Jerman	6:49	✓	✓	✓	✓
RUSAVQAcwuo	Spanyol	7:45	✓	✓	✓	✓
1HB8k17u-5k	Italia	8:07	✓	✓	✓	✓
TFlVWf8JX4A	Inggris UK	9:41	✓	✓	✓	✓
mdulzEfQXDE	Inggris UK	9:56	✓	✓	✓	✓
ru032Mfsfig	Inggris US	10:39	✓	✓	✓	✓
H8lVu0NXplg	Prancis	10:40	✓	✓	✓	✓
xSuNx6xBVWg	Spanyol	11:57	✓	✓	✓	✓
nuMPRqRrgiw	Rusia	14:31	✓	✓	✓	✓
bHlhgxav9LY	Inggris US	14:48	✓	✓	✓	✓
fZEib6B4EOU	Jerman	14:51	✓	✓	✓	✓
iuBTHYlF_Ls	Inggris AU	16:01	✓	✓	✓	✓

di-Tn3zL0NA	Prancis	16:40	✓	✓	✓	✓
cRGJhn2Hv8s	Thai	18:21	✓	✓	✓	✓
gCiWMSl5UGk	Belanda	19:25	✓	✓	✓	✓
2No_CMrxBe8	Inggris US	21:18	✓	✓	✓	✓
6kYT_5eMD7g	Inggris AU	21:48	✓	✓	✓	✓
094y1Z2wpJg	Inggris US	22:08	✓	✓	✓	✓
cUzklzVXJwo	Inggris US	23:29	✓	✓	✓	✓
ZvpGNOWYHa4	Inggris Austria	28:33	✓	✓	✓	✓
Af_6RMIuGSQ	Inggris US	48:21	✓	✓	✓	✓
sJG-rXBbmCc	Inggris US	1:01:25	✓	✓	✓	✓
CaELQS5kTso	Inggris US	1:12:04	✓	✓	✓	✓
CRZhAonfTZs	Inggris Italia	1:24:23	✓	✓	✓	✓
3iEvKW5FJVw	Vietnam	1:24:57	✗	✗	✗	✗
DXHgBUMnlvY	Jepang	1:27:16	✓	✓	✓	✓
K6jgiH2aFhU	Inggris Afrika	1:27:26	✓	✓	✓	✓
0CrKMemAK0g	Inggris UK	1:27:30	✓	✓	✓	✓
p0gnPAnXrqM	Kazakhstan	1:30.00	✓	✓	✓	✓
Cy8KMn3tyPg	Rusia	1:30:27	✓	✓	✓	✓
yoUZZ0zRWmE	Arab Lebanon	1:35:38	✗	✗	✗	✗
hbmf0bB38h0	Inggris US	1:38:28	✓	✓	✓	✓
MzUEZ5F6Htk	Rusia	1:39:49	✓	✓	✓	✓
7c3WQ8gF0QI	Turki	1:40:58	✓	✓	✓	✓
iPJqFVa7lPY	Turki	1:43:52	✓	✓	✓	✓
5dZ_lvDgevk	Inggris US	1:54:16	✗	✗	✗	✗
25YFt6yV-Pg	Hongaria	2:02:33	✓	✓	✓	✓
Frb5hl4aoo4	Hindi	2:29:35	✓	✓	✓	✓
tSbweXvqOhc	Inggris Afrika	3:01:49	✓	✓	✓	✓
LwCRRUa8yTU	Inggris US	6:43:46	✓	✓	✓	✓
HfACrKJ_Y2w	Inggris US	11:53:47	✓	✓	✓	✓
0txtLh8SIOc	Inggris US	11:56:38	✓	✓	✓	✓

This program has been tested to translate various videos on YouTube, from languages in Europe, America, to Asia with long durations ranging from 3 minutes to 12 hours. A total of 52 random video samples from various languages, durations, and content were entered into the program and the results showed that 90.38% of videos were successfully translated into both text and voice, and 9.62% of videos failed to translate because the owner limited interaction. public. Tests carried out on the Application Programming Interface (API) translator program showed that this program has no limitations, only a few conditions are needed for its use, namely the video link must be sourced from YouTube and make sure the internet connection is stable because the program is made online and automatically. In the test, there were several videos that could not be detected in a high-level programming language (python) because the videos had a personal license (license) or copy right (copyright) that could not be published, and high-level machines could not detect objects that which requires further development.
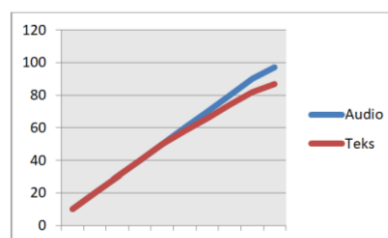

Figure 2. Results of Text and Audio Synchronization Test

Figure 2. shows that in the video synchronization test, the program is able to synchronize text and video. The accuracy between text and voice has been tested and the resulting synchronization value is 89%-97%. To see the results of video samples tested through the Application Programming Interface (API) translator program, you can visit the provided link https://www.youtube.com/watch?v=zasbbNMRh8I.

CONCLUSION

The idea of making a foreign language translator program by utilizing the Application Programming Interface (API) is a new idea as well as an innovation in information techniques due to the limitations of the YouTube platform and video makers providing translations in the form of separate text and audio. In addition, there has not been much use of the Application Programming Interface (API) in linguistics, but many developments have been carried out in other ways using Google Maps, Google Assistant to Google Calendar. With the lack of depth in the field of linguistics, it becomes a challenge to develop linguistic-based programs or software that can help the general public, especially those who have limitations.

The test results show that the program is able to translate the video into text and audio with 90.38% presentation and 89%-97% synchronization with the condition that it must be connected to a stable internet connection. This research has answered the underlying problem. Lack of understanding of a language and other limitations will not be a barrier to obtaining useful information and knowledge. To help you get started, apply the Application Programming Interface (API) to any program that comes with the best software available on the Web.

REFERENCES

[1] Yusri and Mantasiah R, *Linguistik Mikro Kajian Internal Bahasa dan Penerapannya*. Sleman: Deepublish, 2020.

[2] F. Unsiah and R. Yuliati, *Pengantar Ilmu linguistik*. Malang: UB Press, 2018.

[3] N. P. Putri, "Eksistensi Bahasa Indonesia pada Generasi Millennial," *Jurnal Ilmiah Pembelajaran Bahasa dan Sastra*, vol. 5, no. 1, 2017.

[4] D. Eberhard, G. F. Simons, and C. Fenning, *Ethnologue: Languages of The World, 22nd Edition.* Dallas: SIL International, 2019.

[5] I. D. P. Wijana, "Pemertahanan dan Pengembangan Bahasa Indonesia," *Jurnal Ilmiah Kebahasaan dan Kesastraan*, vol. 46, no. 1, 2018.

[6] I. Sari, "Kesulitan Mahasiswa dalam Pembelajaran Bahasa Inggris," *Jurnal Manajemen Tools*, vol. 11, no. 1, 2019.

[7] Irian and Y. Yudhistira, "Implementasi Application Programming Interface (API) Kawal Corona sebagai Media Informasi Pandemik Covid-19 Berbasis Android," *Jurnal Sistem Informasi dan Teknologi Peradaban (JSITP)*, vol. 2, no. 1, 2020.

[8] M. F. A, Muri, H. S. Utomo, and R. Sayyidati, "Search Engine Get Application Programming Interface," *Jurnal Sains dan Informatika*, vol. 5, no. 2, 2019.

[9] K. Yudhistiro, "Pemanfaatan Teknologi Text-to-Speech sebagai Media Pembelajaran pada Laboratorium Bahasa Inggris," *Jurnal Teknologi dan Manajemen Informatika*, vol. 2, no. 1, 2016.

[10] R. G. Guntara, A. Nuryadin, and B. Hartanto, "Pemanfaatan Google Speech to Text untuk Aplikasi Pembelajaran Kamus Bahasa Sunda pada Platform Mobile Android," *Jurnal Sains dan Teknologi*, vol. 4, no. 1, pp. 10-19, 2021.

[11] V. S. Ginting, Kusrini, and E. Taufiq. "Implementasi Algoritma C4.5 untuk Memprediksi Keterlambatan Pembayaran Sumbangan Pembangunan Pendidikan Sekolah Menggunakan Python," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 10, no. 1, 2020.

[12] J. D. Kalleher, *Deep Learning*. Cambridge: The MIT Press, 2019.

[13] P. Goyal, S. Pandey, and K. Jain, *Deep Learning for Natural Language Processing Creating Neural Networks with Python*. Bangalore: Apress, 2018.

[14] M. Hagiwara, *Real-World Natural Language Processing*. Shelter Island: Manning Publication Co., 2021.