UTILIZATION OF GOOGLE CLOUD SPEECH API AND ALCHEMY API ON JOURNALIST INTERVIEW RECORDER ANDROID BASED APPLICATION

Fadli Burhan Hawari[1], Ir. Taryana Suryana, M.Kom.[2]

[1,2] Program Studi Teknik Informatika, Universitas Komputer Indonesia
Jl. Dipatiukur 112-114 Bandung
fadligates.tkj744@gmail.com[1], taryanarx@email.unikom.ac.id[2]

ABSTRACT

Journalistic interview technique is a way to trace information from interviewees to get basic information or complete existing information, one of the tools used in the interview process is by using a voice recorder. Problems that exist in journalism work is composing news, journalist usually takes a long time to compile the news because they have to listen to the recording files of the interviews repeatedly to make text transcription of the recording After in the form of text interviews, journalists have difficulty in compiling the essence of the interview text because they have to sort out the words according to the issues to be raised This certainly requires a long time to process the results of the interview into news. From these problems there is an idea to build an android-based interview recorder application. The technology used in applications that are going to be build is Google Cloud Speech which is used to convert voice recordings into text. Another technology used is Alchemy API which is used to analyze the results of interview text. Furthermore, there are features added to anticipate if journalists interview foreign sources who do not understand each other's languages, namely by utilizing Google Translate technology and Google Cloud Text to Speech which can help applications to translate the interview text into the destination language and then say the text.

Keywords: Android , Google Cloud Speech API, Alchemy API, Google Translate API.

1. INTRODUCTION

News is everything that happens in a condition, then published to the general public and can be used as an appeal, the news itself is presented in electronic or printed form, so that people who are far from the event or condition can know what is happening. And information is data that has been processed into another form that is more useful, namely knowledge or information intended for the recipient in decision making. There are many ways to get information but the most effective way to get a good information is by doing an interview[1]. Journalistic interview technique is a way to trace information from interviewees to get basic information or complete existing information before, one of the tools used in the interview process is by using a voice recorder. After recording the interview, the journalist transcribes the recording into text form by record manual, then the results of the interview transcripts are analyzed to make it easier in the process of drafting a news.

Based on an interview with Intan Silvia Dewi as one of the journalists kumparan there was a problem in the recording media where with many of the files recorded by interviewing Intan Silvia Dewi had to listen to the recorded files one by one stored and record the recording. Another obstacle is the difficulty in compiling the digest from interview transcripts because they have to sort words according to the issues to be raised. This certainly requires a long time to process the results of the interview into a news.

From the above problems researchers have the idea to help journalists to make their work faster and more efficient by speeding up the process of loading news that utilizes technological progress. One of the journalists Intan Silvia Dewi welcomed this idea because it would greatly help the work of a journalist as well as get time efficiency in her work.

The application system that will be built is an application "Utilization Of Google Cloud Speech API And Alchemy API On Journalist Interview Recorder Android Based Application". This application is useful to facilitate the work of journalists in processing the news that will be appointed. Google *cloud speech* API used to convert sound recordings into text and Alchemy API is used to analyze text. In addition to facilitate journalists researchers add speech translator feature that serves to translate one

language into another language which is very useful if the journalist interviews foreign sources who do not understand each other's language.

The purpose of this research is to build an interview recorder application to facilitate the work of journalists using Google Cloud Speech API and Alchemy API technology on the Android platform. While the goal will be achieved in this research are as follows:

1. Make it easier for journalists to compile news without having to listen repeatedly to the results of interview recordings.
2. Make it easier for journalists to draw up the gist of the interview transcript result so as to more easily and quickly in compiling news.

2. RESEARCH CONTENT

2.1 Theoritical Basis

The foundation of this theory explains the concepts that have been systematically assembled that are used to support this research.

2.1.1 Google Cloud Speech API

Google cloud speech API was launched in 2008 in the United States for several types of smartphones. Google's cloud speech API is a framework developed by Google to recognize the voice, converting it to a string (text)[2].

Google's cloud speech API used in this application to convert from an audio file recorded interviews into text form.

2.1.2 Alchemy API

The IBM Alchemy API Cloud Service allows Client to create smart applications that run content analysis on text[3].

Alchemy API is also used to classify content from a text, or to see what topics are trending in the news.

2.1.3 Google Translate API

Google Translate is one of the most famous and most widely used online language translations in the world today, many benefits of Google Translate services to translate languages into other languages easily without having to open a dictionary or other.

Google Translate is an application service provided by Google.Inc, which serves to help translate a text from one language to another[4].

Google Translate API is used in this application to translate text into the destination language so as to facilitate journalists when interviewing foreign sources.

2.1.4 Google Cloud Text to Speech API

Google Cloud Text-to-Speech is a flexible solution to convert text into audio that can be consumed as an audio[5].

Google Cloud Text-to-Speech is used in this application to pronounce translated text into the destination language.

2.2 Analysis and Design

Analysis and design can be interpreted as research on an existing system with the aim of designing a new or updated system.

2.2.1 Analysis of the Problem

Problem analysis is a description of the problem based on the identification of problems in research on the development of a journalist interview recorder application on the android platform.

With the existence of interview activities conducted by journalists with the resource person through the default application on the smartphone there are several obstacles that journalists have difficulty and require a long time to compile a news because they have to listen repeatedly to the audio recordings of the interviews to then be recorded again or used as a transcript, in addition there are other obstacles, namely the difficulty in compiling the digest from the interview transcript because they have to sort out the word according to the issue to be raised.

2.2.2 System Analysis That Is Running

Interviews are always intended as an effort to get the news, comments, opinions in connection with a matter related to expertise owned by a person or resource person who is directly involved with an event or event, because in essence the news is an objective fact that is or has happened.

In conducting interviews, journalists generally use built-in recording media on smartphones. Journalists record during the interview process using a smartphone, then the recording results are heard repeatedly to be used as an interview transcript.

Then after in the form of an interview transcript, the journalist prepares the digest by sorting words according to the issues to be raised. Journalists separating words with the “angle” (angle of view) anything that can later be processed into a news.

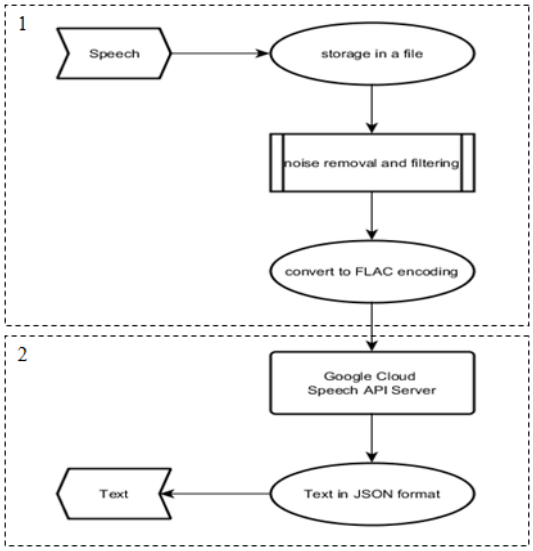
The following is the process flow of journalists in compiling news :

Figure 1. The Flow Of Compilation Of News

2.2.3 Analysis Of Input Data

Input data used in the system to be built is an audio file. In this application researchers use an extension audio file or .flac format. the reason for using the .flac format is because later this audio file will be processed using the Google Cloud Speech API, where Google Cloud Speech only supports audio files with .flac format to convert audio files into text.

Here is the flow of input data from an audio file to become a text:



2.2.4 System Analysis Built

System analysis built is a complete ilustration of the system that will be built. The processes in the system that are built are as follows :

1. The translation process from audio recordings into text form.

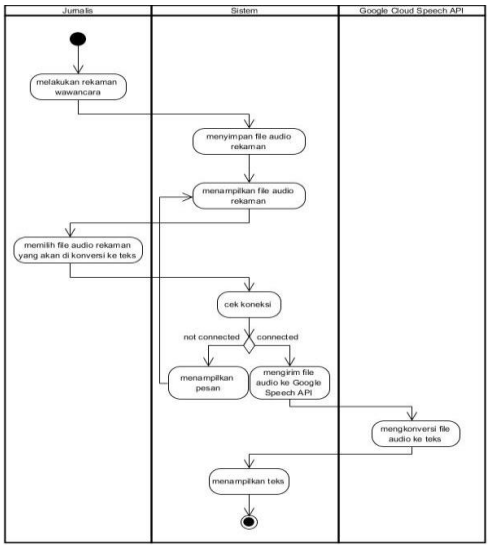


Figure 2. Flow of the System Built Process 1

The following is a description of the flow diagram above:

a. Journalists record interviews with resource persons.
b. The system stores the audio file of the interview recording.
c. The system displays the audio files recorded interviews.
d. Journalists choose the recorded audio file to be converted into text.
e. The system checks the internet connection, if not connected it will display the message and return to the recording audio file display page, if connected, the system will request data to the Google Cloud Speech API.
f. The Google Cloud Speech API converts audio files to text, then the text is sent to the application.
g. The system displays the converted text and saves text to the database.

2. The process of analyzing the text from the results of the translation in the previous process.

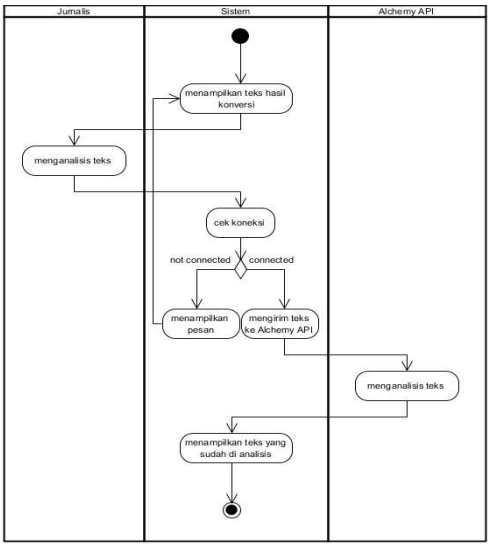


Figure 3. Flow of the System Built Process 2

The following is a description of the flow diagram above:

a. The system displays the converted text in the previous process.
b. The journalist analyzes the converted text.
c. The system checks the internet connection, if not connected it will display the message and return to the converted text page, if connected then the system will request data to the Alchemy API.
d. Alchemy API processes conversion result text to be analyzed.
e. The system displays the text that has been analyzed.

3. The process of translating text of the translation into the target language.

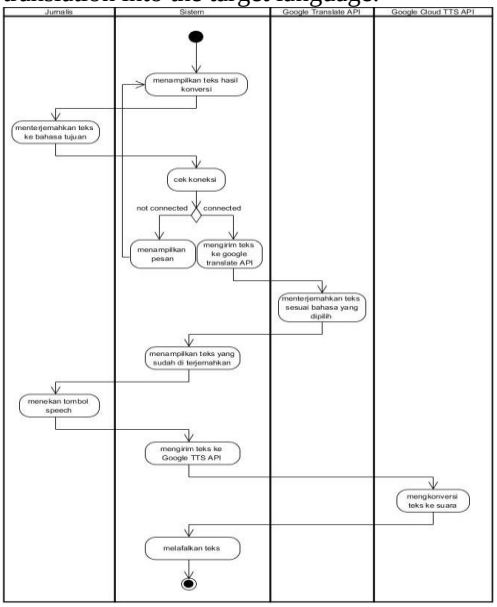


Figure 4. Flow of the System Built Process 3

The following is a description of the flow diagram above:

a. The system displays the converted text in the previous process.
b. Journalists translate the converted text into the destination language.
c. The system checks the internet connection, if not connected it will display the message and return to the converted text page, if connected then the system will request data to the Google Translate API.
d. The Google Translate API translates text according to the language selected.
e. The system displays translated text.
f. Journalist presses the speech button to listen to the spoken text.
g. The system sends translate result text to Google Cloud Text to Speech API.
h. Google Cloud Text to Speech converts text to voice.
i. The system speaks the translated text to the destination language.

2.2.5 System Architecture Analysis

System architecture analysis aims to identify the architecture to be built. The following is the overall system architecture in the following Figure:

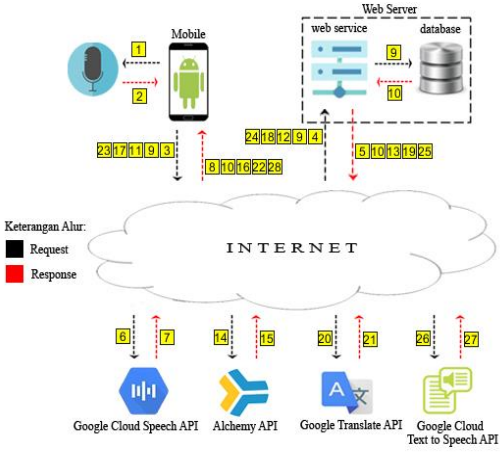


Figure 5. System Architecture

The following is a description of the system architecture built:

1. Journalists make voice recordings.
2. Record files stored on mobile.
3. The application requests to the web service to send audio files to Google Cloud Speech using the internet network.
4. *The Web Service gets a request in the form of a recorded audio file.*

5. Web service response by sending recorded audio to Google Cloud Speech using the internet network.
6. Google cloud speech receives recorded audio files and make the process of speech to text.
7. Google cloud speech API responds in the form of text to mobile via the internet network.
8. The application displays text that has been converted.
9. The application requests to the web service to save audio and text files to the database.
10. The database responds to mobile, audio and text data has been saved.
11. The application requests to the web service to send text files to Alchemy API using the internet network.
12. The Web Service gets a text request.
13. Web service response by sending an interview text to the alchemy API using the internet network.
14. Alchemy API accepts text interview files and processes the text analysis.
15. Alchemy API responds in the form of text analysis to mobile via the internet network.
16. The application displays text that has been analyzed.
17. The application requests to the web service to send text files to the Google Translate API using the internet network.
18. The Web Service gets a text request.
19. Web service response by sending an interview text to the Google Translate API using the internet network.
20. The Google Translate API receives interview text files and make the process of translation into the target language.
21. The Google Translate API responds in the form of text translation to mobile via the internet network.
22. The application displays translated text.
23. The application requests to the web service to send text files to the Google Cloud Text to Speech API using the internet network.
24. The Web Service gets a text request.
25. Web service response by sending translated text, to the Google Cloud Text to Speech API using the internet network.
26. Google Cloud Text to Speech API accepts text interview files and processes the text to speech.
27. Google Cloud Text to Speech API responds to audio files to mobile via the internet network.
28. Application recite the text.

2.2.6 Functional Needs Analysis

Functional needs analysis describes the process of activities that will be applied in a system and explains the needs of the system so that the system can run well and according to needs.

1. Use Case Diagram

   Use Case Diagram provides a way to describe external views of the system and its interactions with the outside world. The following is the use case diagram for the application that was built:

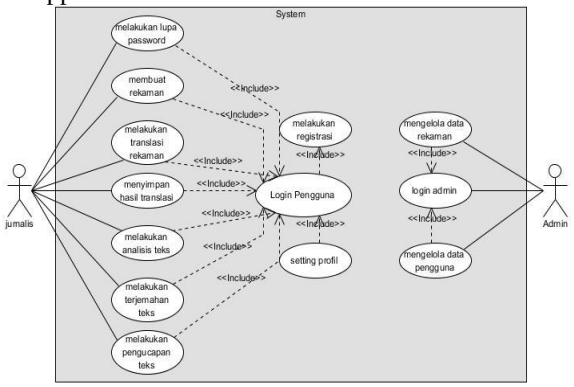


Figure 6. Use Case Diagram

2. *Sequence Diagram*

   Sequence Diagram on the system to be built can be seen in the following figure:

   a. Sequence Diagram Makes Recording

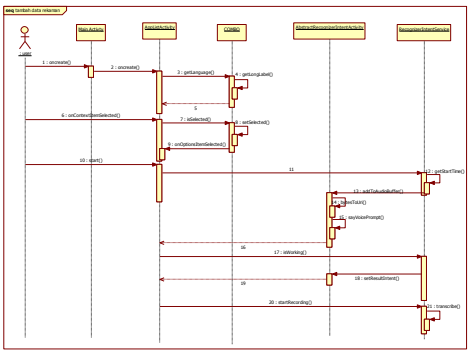


Figure 7. Sequence Diagram Makes Recording

b. Sequence Diagram Performs Record Translation

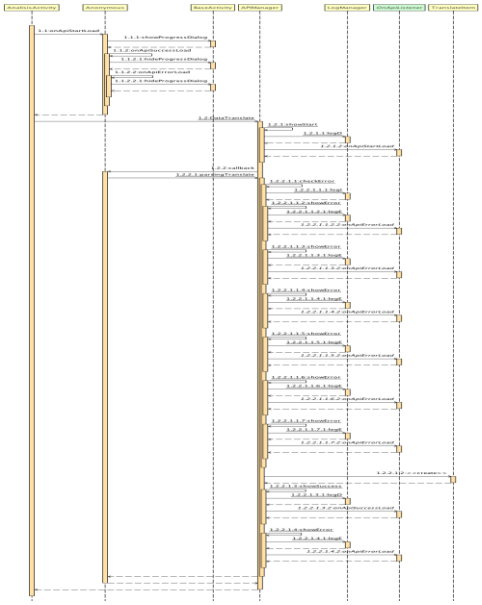


Figure 8. Sequence Diagram Performs Record Translation

2.2.7 Menu Structure

Menu structure is an illustration of the application usage path. The design of the menu structure of this application can be seen in the following Figure:

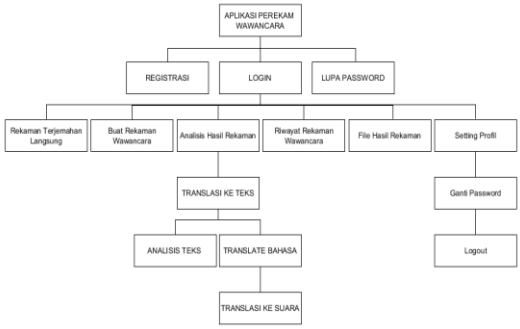


Figure 9. Menu Structure

2.2.8 Implementation Of The Interface

Implementation of the interface is describes the appearance used in the application being built.

Figure 10. Login Interface

Figure 11. Main Page Interface

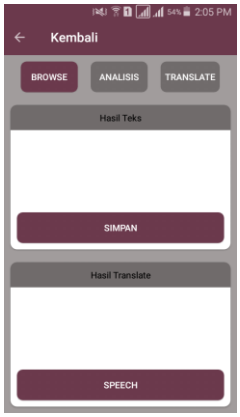


Figure 12. Record Result Analysis Interface

2.3 System Testing

System testing is done to find errors that exist in the system or application that has been built. There are several testing stages: Alpha testing and Beta Testing.

2.3.1 Alpha Testing

Alpha testing is done using the blackbox method that focuses on software functionality whether the software has been running in accordance with the functional needs that have been defined or not.

2.3.2 Beta Testing

Beta testing is a test that takes place in the field to validate the usability, functionality, compatibility and reliability of the system that has been made. This test is done by giving a questionnaire to the users involved in the system being built.

1. Question Number 1
   Does this application make it easier for you to compile news without having to listen to the recording file of repeated interviews ?

Table 1. Results of Question Number 1 Questionnaire

Answer Category	Score	Respondent	Total Score	Value Percentage
Strongly Agree	5	8	40	88%
Agree	4	12	48	
Hesitate	3	0	0	
Disagree	2	0	0	
Strongly Disagree	1	0	0	
Total		20	88	
Hasil	88 / (20x5) x 100% = 88%			

The result of the percentage of respondents is 88% of the expected value is 100%. So it can be concluded, respondents agree that this application makes it easier for journalists in compiling news without having to listen to the recording file of repeated interviews.

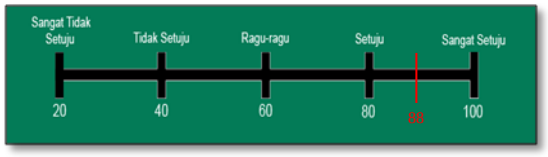


Figure 13. Results of Question Number 1 Questionnaire

2. Question Number 2
   Does this application make it easier for you to compile the digest from the results of the interview transcript ?

Table 2. Results of Question Number 2 Questionnaire

Answer Category	Score	Respondent	Total Score	Value Percentage
Sangat Setuju	5	1	5	69%
Setuju	4	8	32	
Ragu-ragu	3	10	30	
Tidak Setuju	2	1	2	
Sangat Tidak Setuju	1	0	0	
Total		20	69	
Hasil	69 / (20x5) x 100% = 69%			

The result of the percentage of respondents is 69% of the expected value is 100%. So it can be concluded, the respondents almost close to agree that this application makes it easier for journalists to compile the digest from the results of interview transcripts.

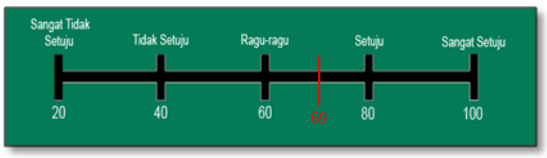


Figure 14. Results of Question Number 2 Questionnaire

3. Question Number 3
   Can the translator feature help you if you meet foreign sources ?

Table 3. Results of Question Number 3 Questionnaire

Answer Category	Score	Respondent	Total Score	Value Percentage
Sangat Setuju	5	4	20	83%
Setuju	4	15	60	
Ragu-ragu	3	1	3	
Tidak Setuju	2	0	0	
Sangat Tidak Setuju	1	0	0	
Total		20	83	
Hasil	83 / (20x5) x 100% = 83%			

The result of the percentage of respondents is 83% of the expected value is 100%. So it can be concluded, respondents agree that the translator feature in this application can help journalists if they meet foreign sources.

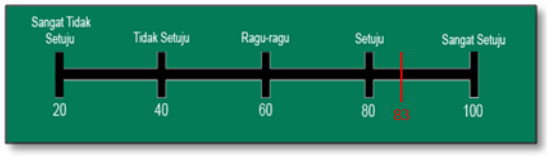


Figure 15. Results of Question Number 3 Questionnaire

3. CLOSING

3.1 Conclusions

After doing the analysis, design, and testing, it can be concluded as follows:

1. This application which is built based on mobile, makes it easy for a journalist to compile a news.
2. This application which is built based on mobile, can accelerate the work process of the journalist in conducting interview transcripts and compile the digest from the interview transcript.
3. With the creation of translate features that utilize the API from Google Translate can facilitate journalists in conducting interviews with foreign sources.

3.2 Suggestions

From the conclusions that have been described above, submitted some suggestions as follows:

1. Make the appearance or interface even more interesting.
2. Develop applications to be used on various platforms.
3. Can translate from audio to text in accordance with the duration of the recording file.

BIBLIOGRAPHY

[1] A. S. Fredy, “Analisis Penggunaan Kata Baku Pada Pengolahan Berita”(Online) https://id.scribd.com/doc/83628336/Analisis-Penggunaan-Kata-Baku. (diakses 08 Juni 2018).

[2] Supriyanta, Pudji Widodo dan Bekti Maryuni Susanto, "Aplikasi Konversi Suara Ke Teks Berbasis Android Menggunakan Google Speech API". Bianglala Informatika Vol 2 No 2 September 2014.

[3] Alchemy API, 2016, "About Application Programming Interface Alchemy" (Online)https://www.alchemyapi.com/about-us. (diakses 10 Mei 2018).

[4] Novita. A. 2013. "Aplikasi Translator Messenger Berbasis Java dan Google Translate API" [Skripsi]. Fakultas Teknik Universitas Negeri Yogyakarta.

[5] Google, "Cloud Text to Speech" (Online) https://cloud.google.com/text-to-speech/. (diakses 10 Mei 2018)