

Organized by

Department of ECE, Adhiyamaan College of Engineering, Hosur, Tamilnadu, India

Chatbot Based Healthcare Monitoring System Using NLP

S.Raghul¹, Dr.S.Suriya²

UG Student, PSG College of Technology, Coimbatore, Tamilnadu, India¹

Professor, PSG College of Technology, Coimbatore, Tamilnadu, India²

ABSTRACT:Text to Speech and Speech to text are some of the developing techniques in the current system. These techniques are used in various domains by Google and other jargons of technology field. Medical field is one such domain which can gain a lot of help from these fields. In current system, people have to check hospitals every time even when they have minor illness. This can be reduced by having a home doctor i.e. a voice bot that can check your illness and giving a suggestion to the user. The voice bot named as “Virtual Doctor” is one such application of speech to text and text to speech. This bot allows the user to give his symptoms via voice or text and analyses them with the existing data on its database and gives back the result in speech and text format. This bot provides a basic medical help provided by a doctor. New symptoms of diseases are often added to the medical database of the bot, so it stays up to date about the diseases.

I.INTRODUCTION

For leading a good life, healthcare is important. Waiting for a doctor’s help at emergency times is so difficult. The proposed idea is to create medical chat bot that helps at these times. It reduces the medical cost. It engages in a voice to voice chat with the patient, diagnoses them and provides medical suggestions.

1.1 TEXT TO SPEECH

Text-to-speech converts text into spoken word, by analysing and processing the text using Natural Language Processing (NLP) and then using Digital Signal Processing (DSP) technology to convert this processed text into synthesized speech representation of the text.

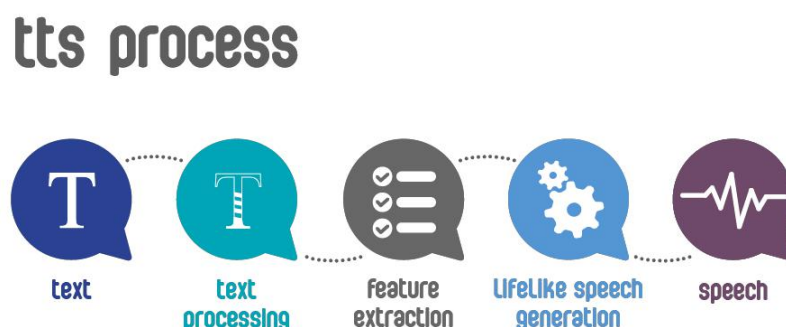


FIG 1.1.1 TEXT TO SPEECH

1.2 SPEECH TO TEXT

The objective of this system is to extract, characterize and recognize the information about speech. Speech to Text records the speech given to microphone as analog signal which is converted into a digital signal. In the converted digital signal, there will be small breaks which indicates the end of one word and beginning of another. Finally, it creates the text.



FIG 1.1.2 SPEECH TO TEXT

II. RELATED WORK

2.1 TEXT TO SPEECH

A Text-to-speech synthesizer is an application that converts text into spoken word, by analysing and processing the text using Natural Language Processing (NLP) and then using Digital Signal Processing (DSP) technology to convert this processed text into synthesized speech representation of the text. Text to speech has two ends. Front end involves text pre-processing and text to-phoneme. The back-end—often referred to as the synthesizer—then converts the symbolic linguistic representation into sound. Concative synthesis is used as it produces more natural language sound. [1]

This paper focuses on converting international language English to text to speech sign. The exchange of text to speech is made by the speech synthesizer which is an imitation technique of human speech. This paper tells about the different ways to perform speech synthesis such as Concentrative synthesis, Articulatory synthesis, Sinusoidal synthesis, etc. The advantages of the system implemented in this paper is that the result achieved from the conversion is audible and perfect. The disadvantage of this system is it only converts only English language. [2]

Complete text is converted to speech segmented to syllables and reading them according to phonetics. After giving the input as text, we have option to select the languages which are available. It is more user friendly and useful application. As it can give its output only as a pure native language, it will be difficult for some people. [3]

Customer support dialog systems Interactive voice response (IVR) systems etc. and are also useful in an applied research. This application is more helpful in banking, toys and many other applications like checking marks, railways, aid to the physically challenged persons, language education and fundamental and applied research. etc. Optical character recognition is used, it is the mechanical or electronic translation of images of and written, type written or printed text. It uses concatenative synthesis, which is a drawback because it requires a pre-recorded database consisting of all frequently used words. [4]

2.2 SPEECH TO TEXT

The objective of this system is to extract, characterize and recognize the information about speech. This system is implemented using Mel-Frequency Cepstral Coefficient (MFCC) feature extraction technique and Minimum Distance Classifier (MDC), Support Vector Machine (SVM) methods for speech classification. MFCC is used for finding the difference between human voice and background noise, MDC is used for classification of words (training and testing). Training the system to differentiate the words and external sounds is a very tedious process. [5]

This paper investigates the speech to text conversion of Kannada language. The speech processing is done using CMU Sphinx framework. Language dependent models like Acoustic model which represents possible range of audio representations and Language model which models pattern probability of the words are used for accurate interpretation of words spoken. It makes use of thousands of sentences as test cases to train the model which results in high accuracy of understanding the way the people pronounce the words during speech which lays the major advantage for the system. But this accuracy can be achieved only when the user speaks in a proper pace which is a disadvantage. [6]

Speech, which is represented as acoustic waves is converted to text by recognising the words and converting them into sentences. Conversion that is happening will not be so accurate and is estimated by summarizing words and sentences. This is majorly based on speech extraction and concatenation, which increases the accuracy of conversion. In order to find the speed at which the conversion is happening, words per minute are calculated and are trying to improve the speed. [7]

Linear Predictive Coding to extract important features from voice signal and Neural Network with Back propagation to classify and recognize spoken words. The spoken words are converted to signals and analysed using a set of parameters. These parameters will be trained or compared with the given samples in the database to identify the words. The analog is converted to digital and analysed using linear Predictive Coding. After analysing, speech can be extracted into four important parameters, Voiced/Unvoiced Decision, Pitch Period, Gain Factor, and LPC Coefficient. After LPC, neural network architecture is used for pattern matching. Considering all neurons in the network have been calculated through an activation function, the output pattern is computed. This output is then compared with the target output. Weight assigned will be stored if the resulting output and target output is matched. Otherwise, error is calculated and weight will be adjusted. Accuracy is said to be 95% if enough training set is given. [8]

2.3 CHATBOT

This chatbot has mainly dealt with the patients in detecting and diagnosing the illness caused to patients. It will detect all the health problems that are in knowledge base and will be limited to some domain. It will not only give direct tips but also suggests to meet a specific specialist for the problems. It helps to identify the disease even if we don't know what disease it is. It also got good reviews that it is detecting and helping others with more accuracy. [9]

This paper talks about a chatbot which is a text-to-text bot that acts as a medical reference book. This chatbot engages in conversation with their users about their medical issues and provides a diagnosis based on their symptoms. This bot uses artificial intelligence to predict and diagnose the disease. It also takes in account of the previous chat histories to determine a pattern related to the patients' health. The major advantage of this bot is that it does immediate diagnosis of the patient without making them move from their place. Since it is a medical care bot, the diagnosis make by it are only okay for first level assumption of the diseases and hence the patient has to meet the doctor at last for the cure. [10]

By introducing an artificial brain, the web-based bot generates customized user responses, aligned to the desired character. Questions asked to the bot, which is not understood is further processed using a third-party expert system (an online intelligent research assistant), and the response is archived, improving the artificial

brain capabilities for future generation of responses. Voice recognition requires a two part process of capturing and analysis of an input signal. Server response generation can be broken down into two categories: data retrieval and information output. The output i.e. the response of the bot should match the user’s query. The web service processes all received queries using the response generation module (based on the Artificial Linguistic Internet Computing Entity (ALICE) [10] system), which makes use of a data repository and respond to the user. The data repository is updated by the content retrieval module to increase the intelligence autonomously (based on an Artificial Intelligence Markup Language (AIML)). The intelligence is limited. [11]

The proposed Medical Chatbot can interact with the users, giving them a realistic experience of chatting with a Medical Professional. It can detect human message patterns using AIML (Artificial Intelligence Mark-up Language) which is a mark-up language based on XML to build AI applications. It retrieves keywords from the initial messages to know the possible medical problems that the user has, based on their input. It also provides Recommended medication, remedy, while the Health Tap (Facebook Messenger Chatbot) only mentions previously answered questions by doctors. Not mentioned any training techniques, so it can’t answer for any new disease or any disease which is not trained into the system. [12]

III.ANALYSIS AND DESIGN

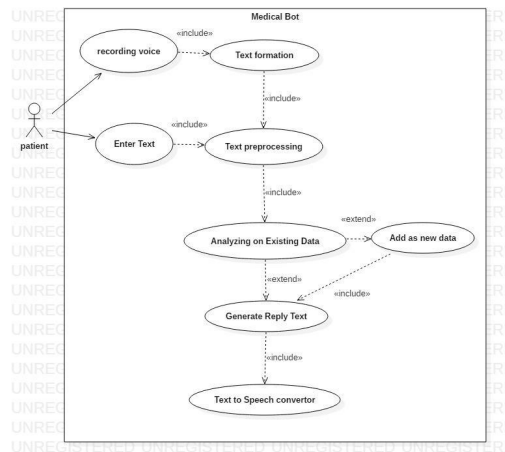


FIG:3.1.3 USECASE DIAGRAM

Use case Functionality	Description
Recording Voice	When user speaks, his voice is recorded
Text Formation	User's voice message is converted to a text format.
Text Pre-processing	Irrelevant information from the converted text is pre-processed i.e. removed. Only relevant text is used in further functions.
Analysing data	The pre-processed text is analysed in the existing trained data and a corresponding reply text is generated
Adding data	If the text is not found in trained data, then it is added to the training set
Reply Text	After analysing the data , a reply text is generated.
Text to speech convertor	The generated reply is converted to speech and played to the user.

Table 3.1.1 Use cases Description

IV.SYSTEM IMPLEMENTATION

4.1 Dialogflow

Dialogflow is a platform designed specifically for human-computer interaction. The company formerly called as API.ai was bought by Google and was later renamed as Dialogflow. The platform is known for developing chatbots and assistances that performs certain tasks and answers to user's questions. Dialogflow has an in-built natural language processing engine that was later redesigned by google for better interaction experience. In Dialogflow, there are many components that are used for developing an interactive service and let's take a look at each of those below.

4.2 Agent

In all the components present in Dialogflow, the basic component is the agent. Agents are best portrayed as Natural Language Understanding (NLU) modules. These modules can be incorporated into your application and decipher message or talked client demands into significant information. This interpretation happens when a client's spoken words or texts matches an intent inside the agent. There can be any number of intents with an agent.

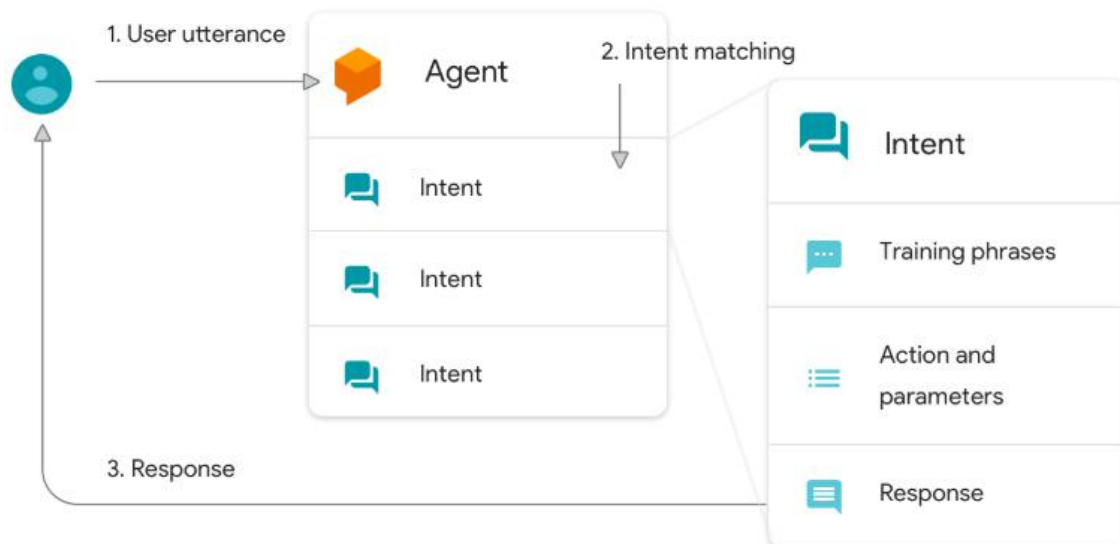


FIG 4.2.4 SYSTEM IMPLEMENTATION

The above image is an example of how the user's utterances matches with an intent in an agent. Whenever a user speaks or text, its checks all of the intents in the agent. The intent matches will be triggered to give responses. Dialogflow underpins many root dialects that can be utilized for the fundamental advancement of the agent. Some root dialects incorporate regions, which cover area or nation particular variants of a root dialect. It also uses machine learning algorithms for better understanding of natural language utterances. In virtual doctor, the agent used is called Cold-Agent.

4.3 Intent

Another major component of this platform is called an intent. Intent defines the conversation between the user and the agent. In intents we can define the examples of what user's may speak and what responses should be given for the defined user's utterances. If two or more intents have similar triggering conversations, the one that matches the most will be chosen for responses. Intent can also be triggered in case of events. We can specifically tell that a intent should be triggered when events such as Welcome or Google Assistant Welcome occurs. The 5 types of intents are chest pain intent, cold intent, fever intent, headache intent and Sore throat intent.

4.4 Entities

Entities are used for identifying and extracting useful data from the user's inputs. While intents are used to match the users input, entities allow us to extract data from user's utterances which can be stored as variables and used later in responses. There are many default entities like dates, time, days, etc. We can also customise entities according to the user's needs.

In virtual doctor, the entities defined are cold, headache and sorethroat. Inside these entities we have to define what words or sentences should be matched to that particular entity. To understand how this works, let look into an example. The following is the image of defined terms inside duration entity.

4.5 Contexts

Contexts represent the current state of the user’s request and allow the agents to carry information for one intent to another. It is the one that helps us to control the conversational flow between the users and the agent. There are generally two types of contexts namely, input and output contexts. If the input context is applied to an intent, it tells the agent to match the intent only if the user’s utterance is a close match and also if the context is alive. The output intent if applied to an intent, makes the context active and also maintains the context after the intent is matched. The lifespan of a context is, it either expires either after 20 minutes after its match with the corresponding intent or after five requests.

4.6 Training

Training component helps us build and refine our agents using real world data. We can also make use of existing sources of conversation data, as well as usage and performance data that pertains to your running agents.

There are three main use cases for real world data within Dialogflow:

1. Building a new Dialogflow agent using logs of existing customer interactions.
 - Use training to add existing data to training phrases.
2. Improving the performance of a live Dialogflow agent using its own logs.
 - Use training to add additional data to training phrases.
3. Understanding the performance of a live Dialogflow agent to inform your design decisions.
 - Use analytics to assess performance of agent.

Training also enables us to manually train the data. Usually, a history of conversation between the user and the agent will be recorded and this data can be used for training. This can be used to assign unmatched phrases to match an intent. For example, let’s take a look at the image of training data below.

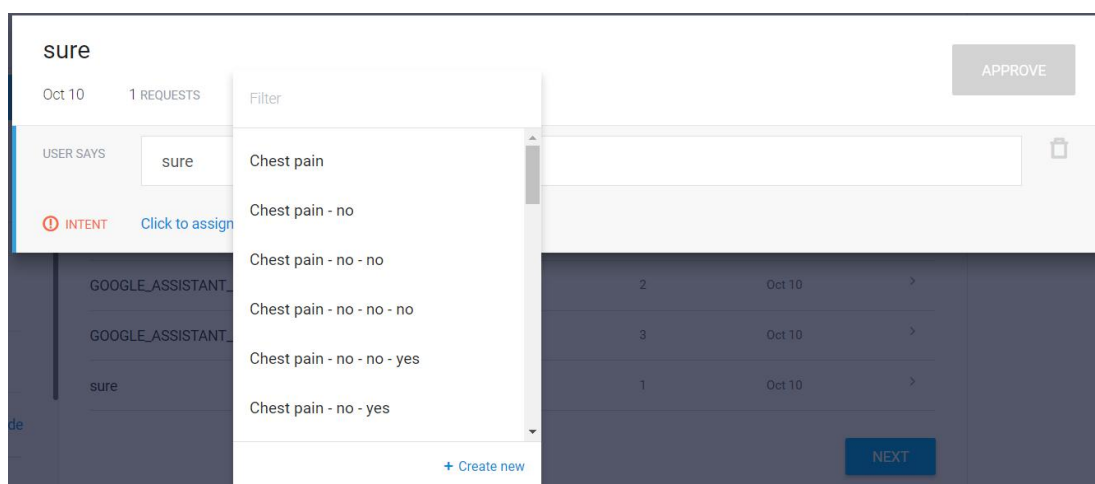


Fig 4.6.5 Training

In the above image, “sure” is the word uttered by the user. But it does not match with any of the intents and throws error. But in general, say it actually matches an intent and we didn’t train it properly. In that case, we can check the training data, and assign the word to an intent manually. After assigning, whenever the user utters the same word or phrase again, it will trigger the assigned intent.

V. TESTING

This is the most interesting part of Dialogflow. Since google has bought it, the developers in this platform has a major advantage and also a great privilege of integrating with many popular social platforms like Google Assistant, Facebook Messenger, etc. This was also one of the main reasons that made us to choose Dialogflow. To make use of these platforms, the application developed has to go through various procedures and tests conducted by the platform that we choose to use. These tests and procedures may vary from one platform to another. After going through these test, the company owing that platform will send the result. If the application passes all the tests, it is eligible for the production i.e., the application can be used by all the people using that platform. The following are the list of all the platforms supported by Dialogflow.

- Dialogflow Web Demo
- Google Assistant
- Facebook Messenger
- Dialogflow Phone Gateway
- Slack
- Viber
- Twitter
- Twilio IP
- Twilio (Text Messaging)
- Skype
- Tropo (Text Messaging)
- Telegram
- Kik
- LINE
- Cisco Spark
- Amazon Alexa
- Microsoft Cortana

In alpha testing, there can only be 20 users or less. The application should be tested by the users for at least 10 days. If so, then the application is ready for beta testing. Virtual doctor went through rigorous alpha testing and hence was eligible for beta testing.

For beta testing, there is a checklist rules to be followed like checking for violation policies, following the brand guidance, testing our actions in the console, registering the actions done, etc. There can be up to 200 users in beta testing. Only if all the application satisfies all the rules in the check list, it will be eligible for production. The virtual doctor satisfied all the rules in the checklist only after its fourth time and is now eligible of production. The following is the image of the e-mail sent from Google regarding the result of beta testing of the virtual doctor.

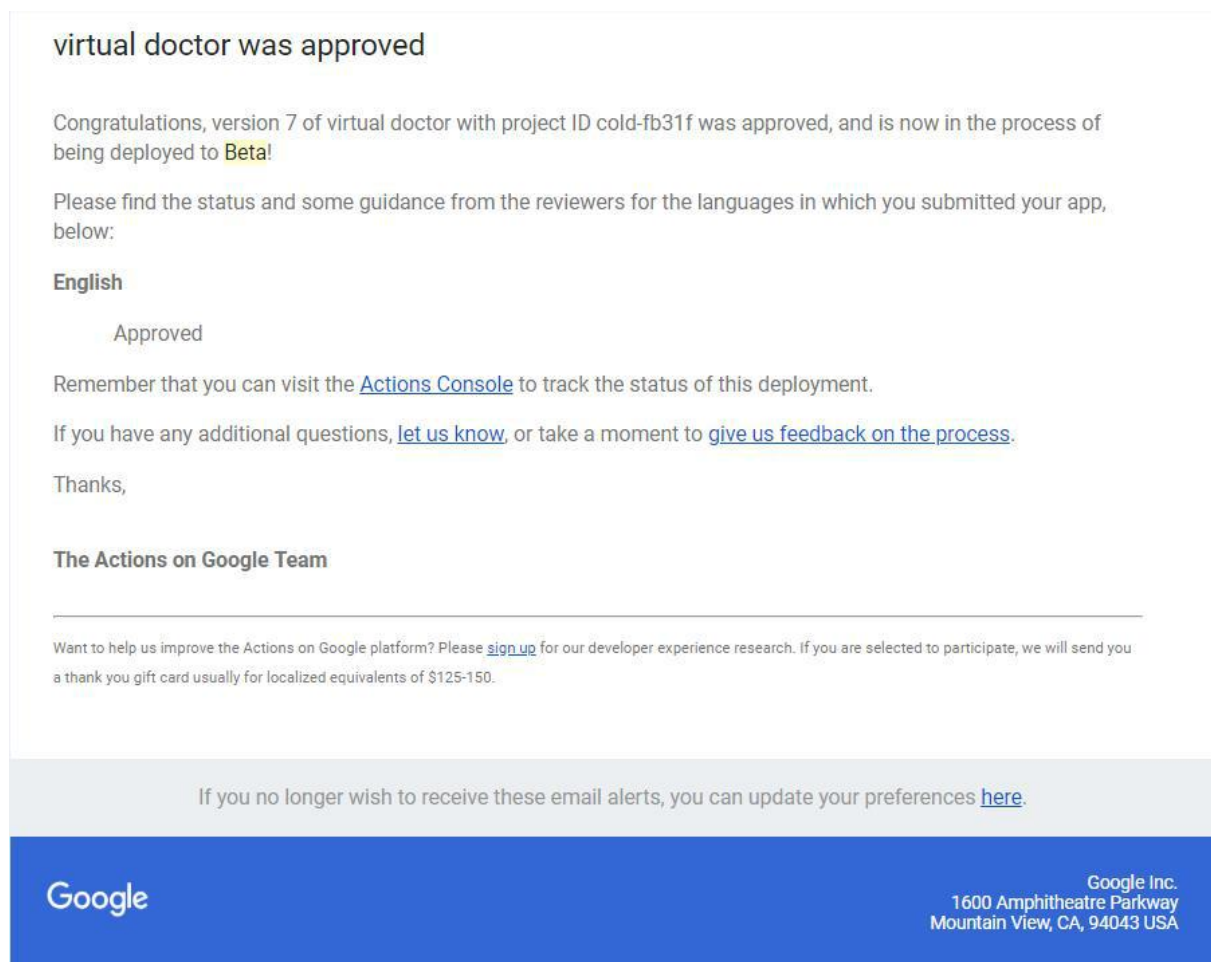


Fig 5.1.7 Testing

VI. CONCLUSION AND FUTURE SCOPE

Virtual Doctor will be really help for the people suffering with some common diseases and it will give the required medication based on some symptoms. This will be very useful for any kind of people who knows to talk English because it can be completely dealt with voice. It is very useful for people who will not have proper hospitals near their locality.

Organized by

Department of ECE, Adhiyamaan College of Engineering, Hosur, Tamilnadu, India

In future, we want to extend Virtual Doctor for multi-language support and we want to add medication for some more diseases such that for each and every disease, we want to use Virtual Doctor.

REFERENCES

- [1] <https://www.ijais.org/archives/volume7/number2/623-1143>
- [2] http://pep.ijeee.org.in/journal_pdf/11-74-140689395029-35.pdf
- [3] <https://acadpubl.eu/jsi/2017-115-6-7/articles/6/53.pdf>
- [4] http://ijarcse.com/Before_August_2017/docs/papers/Volume_5/10_October2015/V5I10-0228.pdf
- [5] <https://ieeexplore.ieee.org/document/7754130/>
- [6] <https://ieeexplore.ieee.org/document/7830119/>
- [7] <https://ieeexplore.ieee.org/document/1306513/>
- [8] <http://ieeexplore.ieee.org/document/5072010/>
- [9] <https://ieeexplore.ieee.org/document/8280270/>
- [10] <http://matjournals.in/index.php/JoWDWD/article/download/2334/1613>
- [11] <https://ieeexplore.ieee.org/document/5167660/>
- [12] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3090881
- [13] <https://dialogflow.com/>
- [14] <https://tutorials.botsfloor.com/building-your-own-chatbot-using-dialogflow-1b6ca92b3d3f>
- [15] <https://www.drugs.com/symptom/coughs-and-colds-1.html>
- [16] https://www.medicinenet.com/sore_throat_pharyngitis/article.htm

virtual doctor was approved

Congratulations, version 7 of virtual doctor with project ID cold-fb31f was approved, and is now in the process of being deployed to **Beta!**

Please find the status and some guidance from the reviewers for the languages in which you submitted your app, below:

English

Approved

Remember that you can visit the [Actions Console](#) to track the status of this deployment.

If you have any additional questions, [let us know](#), or take a moment to [give us feedback on the process](#).

Thanks,

The Actions on Google Team

Want to help us improve the Actions on Google platform? Please [sign up](#) for our developer experience research. If you are selected to participate, we will send you a thank you gift card usually for localized equivalents of \$125-150.

If you no longer wish to receive these email alerts, you can update your preferences [here](#).

Google

Google Inc.
1600 Amphitheatre Parkway
Mountain View, CA, 94043 USA