

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/345315840>

# Desarrollo de nuevos comandos para Google Assistant utilizando Dialogflow, Firebase y NodeMCU

Article · October 2020

CITATIONS

0

READS

83

3 authors, including:



**Ignacio López Martínez**

Tecnologico nacional de Mexico, campus Orizaba

2 PUBLICATIONS 0 CITATIONS

[SEE PROFILE](#)



**Manuel Suárez Gutiérrez**

Universidad Veracruzana

14 PUBLICATIONS 2 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Análisis de información proveniente de la red social online de Twitter, utilizando técnicas de Big Data [View project](#)

## **Desarrollo de nuevos comandos para Google Assistant utilizando Dialogflow, Firebase y NodeMCU (ESP8266) como intermediario**

### **Development of new commands for Google Assistant using Dialogflow, Firebase and NodeMCU (ESP8266) as an intermediary**

Noé Hernández García <sup>1</sup>, Ignacio López Martínez <sup>2</sup>, Manuel Suárez Gutiérrez<sup>3</sup>

<sup>1</sup> División de Estudios de Posgrado e Investigación, Maestría en Sistemas Computacionales, Instituto Tecnológico de Orizaba, Orizaba, Veracruz, México

<sup>2</sup> División de Estudios de Posgrado e Investigación, Maestría en Sistemas Computacionales, Instituto Tecnológico de Orizaba, Orizaba, Veracruz, México

<sup>3</sup> Instituto de Investigaciones y Estudios Superiores Económicos y Sociales, Universidad Veracruzana, Xalapa, Veracruz, México

noe-hdezg@hotmail.com, ilopezm@orizaba.tecnm.mx, mansuarez@uv.mx

#### **Abstract**

Home automation is an area of great interest and innovation, since its main objective is the automation of various tasks in the home. To achieve a domotic environment, it is necessary the intervention of certain specialized devices, such as virtual assistants, which are nothing more than computerized agents in charge of performing several actions with very little user interaction, such interaction is mostly carried out by voice commands sent by the user to the device, as in the case of Google Home Mini. Thanks to a study related to the subject of virtual assistants, it was concluded that there is very little information regarding the creation of new voice commands for virtual assistants, and the absence of a guide to help in the creation of these. Therefore, in addition to knowing the characteristics of the aforementioned technologies, the present proposal was reached: creation of new voice commands for the Google Assistant. This virtual assistant is integrated in the Google Home Mini, which will be used to test the new voice commands resulting from the use of the development architecture. In addition, the use of Dialogflow and the Google Assistant SDK (Software Development Kit) is proposed.

#### **Resumen**

La domótica es un área de mucho interés e innovación, ya que su objetivo principal es la automatización de diversas tareas en el hogar. Lograr un ambiente domótico, es necesario la intervención de ciertos dispositivos especializados, como los asistentes virtuales, que no son más que agentes computarizados encargados de realizar varias acciones con muy poca interacción de los usuarios, dicha interacción es mayormente llevada a cabo por comandos de voz que manda el usuario al dispositivo, como en el caso del Google Home Mini. Gracias a la realización de un estudio

relacionado con el tema de los asistentes virtuales, se llegó a la conclusión de que existe muy poca información con respecto a la creación de nuevos comandos de voz para asistentes virtuales, y la ausencia de una guía que de paso a la creación de estos. Por lo que además de conocer las características de las tecnologías antes mencionadas, se llegó a la presente propuesta: creación de nuevos comandos de voz para el Google Assistant. Dicho asistente virtual está integrado en el Google Home Mini, mismo que será utilizado para probar los nuevos comandos de voz resultantes de la utilización de la arquitectura de desarrollo, además, se propone la utilización de Dialogflow y el SDK (Software Development Kit, Paquete de Desarrollo de Software) de Google Assistant.

---

Keywords and phrases: Asistente Virtual, Automatización, Domótica.

2020 Mathematics Subject Classification: 68N01.

---

## 1 Introducción

Los asistentes virtuales inteligentes en [1] son un conjunto de programas informáticos capaces de interactuar con los seres humanos en su propio lenguaje, esto quiere decir que son entidades autónomas que ofrecen colaboración a las personas para alcanzar objetivos a través de diferentes canales de interacción. Los asistentes virtuales se pueden clasificar de la siguiente manera: aquellos que tienen un dispositivo físico y los que se encuentran embebidos en los sistemas móviles.

Debido a que no hay un estudio que compare las diferentes tecnologías para el desarrollo e implementación de un conjunto de instrucciones que no están presentes actualmente para el asistente virtual Google, ello conlleva que las personas que busquen una técnica adecuada para el desarrollo de un conjunto instrucciones según sus necesidades. Este trabajo resuelve la problemática sobre el desarrollo de un conjunto instrucciones usando el asistente virtual de Google para el encendido y apagado de diversos dispositivos externos caseros.

En el Marco Teórico, se encuentran los conceptos y definiciones necesarias para entender el alcance de este proyecto, mientras que en la sección de Trabajos Relacionados, se exponen los artículos relacionados al tema, en el Proceso de Desarrollo, se explica directamente el proceso de desarrollo de la metodología propuesta en este documento, misma que será detallada en la sección de Desarrollo de la Arquitectura; finalmente, en la parte de Conclusiones y Recomendaciones, se limitan los alcances de la metodología y, en Trabajos a Futuro, se plantea la aplicación de dicha metodología en un caso de estudio para probar su utilidad y aplicación.

## 2 Marco Teórico

### 2.1 Domótica

La integración tecnológica de los sistemas electrónicos en el hogar se conoce como domótica. En [2] se define este término como agrupaciones automatizadas de equipos, normalmente asociados por funciones, que disponen de la capacidad de comunicarse interactivamente entre ellos a través de una conexión multimedia que los integra.

## 2.2 Lenguaje natural

El lenguaje natural es el conjunto de signos y símbolos orales por medio de los cuales los seres humanos se comunican entre sí, a través de la formación de palabras que expresan ideas [3].

El uso del lenguaje natural en los asistentes virtuales es una pieza clave en el desarrollo de nuevos comandos e instrucciones para el Google Assistant, ya que es la forma en que el usuario interactúa con este último; si bien la complejidad de identificar, más que entender dicho lenguaje está cubierta por el propio asistente virtual, aún existen adecuaciones a la forma de expresarse que no se interpretan de forma correcta

## 2.3 Procesamiento de lenguaje natural

El Procesamiento del Lenguaje Natural (PLN) es una transformación del lenguaje natural a una representación formal, dicha representación se manipula y de ser necesario, sus resultados se llevan nuevamente al lenguaje natural.

El PLN incluye la recuperación y extracción de información, traducción automática, sistemas de búsquedas de respuestas, generación de resúmenes automáticos, minería de datos y análisis de sentimientos [4].

## 2.4 Agentes inteligentes

Un agente es una entidad autónoma capaz de almacenar conocimiento sobre sí misma y sobre su entorno, de otra forma, es un programa que, basándose en su propio conocimiento, realiza un conjunto de operaciones para satisfacer las necesidades de un usuario o de otro programa, bien por iniciativa propia o porque alguno de los anteriores se lo ordena [5]. El agente inteligente de Google es capaz de analizar el lenguaje natural y obtener las palabras clave que sirven para ejecutar una acción o instrucción.

## 2.5 Google Assistant

El asistente virtual personal de Google es una tecnología desarrollada con inteligencia artificial, su principal característica es la conversación bidireccional entre los usuarios y Google, donde las preguntas y respuestas se pueden complementar sin restringirse a un solo tema [6]. Brinda la posibilidad de comunicación bidireccional en forma de una conversación real. El usuario puede interactuar con Google Assistant (GA) con la ayuda de comandos de voz gracias a sus algoritmos de procesamiento de lenguaje natural, además, realizar búsquedas en Internet, programar citas, configurar alarmas y hasta administrar la capa de sesión de medios, también se encuentra como un servicio disponible en la nube.

Además, Google Assistant, brinda posibilidades de incrementar las funciones o acciones que permitan alcanzar un mayor grado de interacción con los usuarios, para ello se dispone de un entorno de desarrollo proporcionado por el propio Google.

## 2.6 SDK de Google Assistant

El SDK de Google Assistant es una tecnología que permite agregar nuevos comandos de voz a través de la detección de palabras clave y mediante la comprensión del lenguaje natural y el uso de la inteligencia de Google para recibir ideas.

El servicio de Google expone una API de bajo nivel que le permite manipular directamente los bytes de audio de una solicitud y respuesta del Asistente. Se pueden generar enlaces para esta API en

lenguajes como Node.js, Go, C ++ y Java para todas las plataformas que admiten gRPC, este un moderno framework de código abierto y alto rendimiento que puede funcionar en cualquier entorno.

Primero se captura una solicitud de audio hablada, ésta se envía al asistente y posteriormente se recibe una respuesta de audio hablada, además del texto en bruto de la emisión [7].

Anteriormente para llevar a cabo una acción deseada, se debía hacer uso de árboles conversacionales, los cuales permitían mostrar la interacción del lenguaje natural con el SDK de Google; estos árboles conversacionales eran transmitidos mediante un formato de texto denominado JSON, el cual se ayudaba de toda la gama de lenguajes de programación que admite gRPC.

### 3 Trabajos Relacionados

En [8] se estableció que el uso de asistentes inteligentes y la automatización de hogares inteligentes se vuelven cada vez más fuerte. Los asistentes virtuales habilitados para voz ofrecen una amplia variedad de servicios orientados a la red y, en algunos casos, logran conectarse a entornos inteligentes, mejorándolos con interfaces de usuario nuevas y efectivas, sin embargo, dichos dispositivos revelan nuevas necesidades y debilidades, por ello, se desarrolló un nuevo asistente inteligente con la capacidad de observar un rostro y determinar las emociones de este; dicho asistente inteligente se ejecuta en un dispositivo Raspberry Pi 3 por medio de una interfaz gráfica desarrollada en HTML5 que contiene JavaScript necesario para comunicarse con una conexión WebSocket y que implementa la API (Application Programming Interface, Interfaz de Programación de Aplicaciones) RESTful con los módulos de síntesis y reconocimiento de voz, esto permite una mejor interacción entre el usuario y el asistente inteligente.

Debido a que los comandos de voz se encuentran expuestos a diversas situaciones que alteran el entendimiento del lenguaje natural en los dispositivos, Michael et al. [9] presentó un nuevo sistema de localización de palabra clave, este sistema que utiliza el ASR (Automatic Speech Recognition, Reconocimiento Automático de Voz) del lado del servidor y una frase desencadenante para mejorar la precisión general de KWS (Knowledge work system, Sistemas de Manejo de Conocimiento), el cual, permite al usuario hablar sin interrupciones entre la frase de activación y el comando de voz, reduciendo así significativamente la tasa de falsas aceptaciones (FA) en un 89%, al mismo tiempo que aumenta de manera mínima la tasa de falsos rechazos (FR) en un 0.2% para mejorar la calidad de ASR.

En [10] C. Peng y R. Chen desarrollaron una función hecha a medida para los usuarios, el cual consiste en el encendido y apagado de un socket, utilizando el reconocimiento de voz de Google Home con la concepción del aprendizaje automático para probar el análisis de viabilidad sobre cómo satisfacer las necesidades de los usuarios mediante un patrón de hogar inteligente con el diseño del aprendizaje automático utilizando LSA (Latent Semantic Analysis, Análisis semántico latente) and TF-IDF (Term Frequency–Inverse Document Frequency, Frecuencia de Término – Frecuencia Inversa de Documento). Para conectar el dispositivo móvil con el Raspberry Pi se utilizó BluePy que es una suite que proporciona API (Application Programming Interface, Interfaz de Programación de Aplicaciones) para conectar Bluetooth bajo consumo basados en el lenguaje Python el cual se conecta con el dispositivo Google Home, el Bluetooth abre una conexión utilizando API.ia hoy conocido como Dialogflow permitiendo conversaciones en lenguaje natural humano-computadora

mediante su SDK, este permitirá la creación de la nueva función. El experimento permite a los usuarios enviar comentarios al reconocimiento de voz de Google Home y luego transferir la señal de Bluetooth a Raspberry Pi para controlar los dispositivos.

Hwang et al. [11] propusieron una arquitectura para la generación automática de guías de interacción de usuario con asistente inteligente. Dado que el mayor obstáculo es la dificultad que tienen los usuarios conocer el alcance del servicio (característica o conocimiento) y, por lo tanto, es difícil entregar la solicitud correcta. Si el asistente inteligente sugiere las funciones adecuadas basadas en el contexto o las preferencias del usuario con el idioma o las voces naturales, podría hacer que los usuarios estén familiarizados con él, por lo que podría hacer que los asistentes inteligentes sean muy utilizados. Para el desarrollo de esta arquitectura se utilizó una ontología denominada red de tareas, esta red de tareas se usa para encontrar la secuencia correcta de acciones para lograr la solicitud del usuario. Se utilizan los métodos NLG/TTS (Natural Language Generation/Text-To-Speech, Generación de Lenguaje Natural/Texto A Voz) para hacer la guía natural basada en los planes que el sistema desea recomendar.

Cheng et al. [12], desarrollaron y evaluaron una aplicación con el asistente de Google Home que actúa como una estrategia de intervención innovadora para ayudar a los pacientes ancianos con el autocontrol de la diabetes tipo 2. Este marco de aplicación combinó la interfaz de voz de Google Home para alojar el agente de conversación utilizando la plataforma de API.ia hoy conocido como Dialogflow en conjunto con el SDK de Google y una interfaz web para la visualización de datos se creó un WebHook, el cual consiste en devolución de llamada HTTP (Hypertext Transfer Protocol, Protocolo de Transferencia de Hipertexto) utilizando Node.js como plataforma, este utiliza JavaScript con lenguaje de programación y con el fin de reducir la carga de monitoreo sobre las consecuencias diabéticas para el usuario. Se realizó una comparación basada en las características de la aplicación con las aplicaciones móviles disponibles actualmente para determinar el cumplimiento relativo de los componentes oficiales. Como resultado la aplicación mejoró el estado actual de la técnica al aumentar la satisfacción y conveniencia del usuario.

En breve, se muestra una tabla comparativa con respecto a las características analizadas en los artículos antes expuestos y los que posee la propuesta del trabajo presente.

Tabla I. Comparación de Trabajos Relacionados

Art.	Características				
	1	2	3	4	5
[8]	X		X		
[9]		X	X		
[10]	X		X	X	
[11]		X			
[12]	X		X		X
Propio	X	X	X	X	X
1) Tecnología de Google, 2) Arquitectura, 3) Nuevo desarrollo 4) Implementación, 5) Evaluación					

Como se observa en la tabla 1, la tecnología de Google se encuentra presente en la mayoría de los trabajos relacionados. La propuesta de este trabajo es mostrar los resultados sobre el desarrollo e implementación de nuevos comandos de voz para el asistente de Google, utilizando la tecnología de Google, aplicando una nueva arquitectura para interconexión con los dispositivos. La literatura actual no ofrece suficiente información de casos relacionados con el desarrollo de nuevos comandos e instrucciones para el asistente virtual de Google, ya que éste no ha sido completamente explotado, debido a que en algunos casos resulta ser difícil de utilizar o de estudiar.

#### 4 Proceso De Desarrollo

Para cumplir con el objetivo de este trabajo, se utilizó la metodología descrita en la figura 1 que se muestra a continuación.

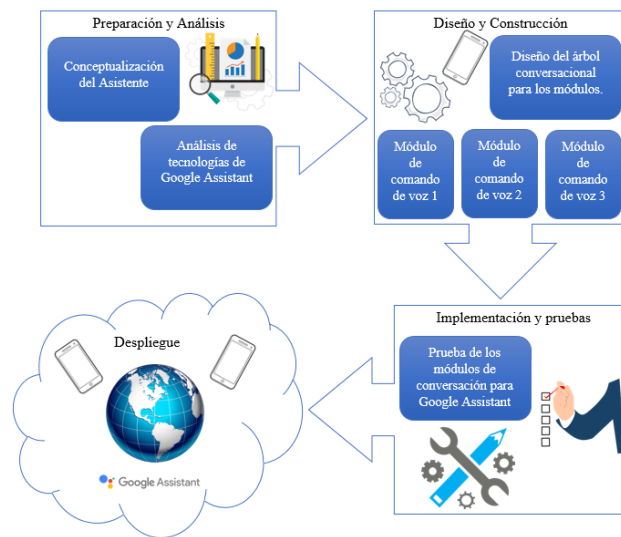


Figura 1. Proceso para el alcance de la solución del proyecto.

- **Preparación y análisis:** En esta etapa, se analizaron diferentes investigaciones disponibles en las bibliotecas digitales de ACM Digital Library, Springer Link, Science Direct y IEEEExplore, con el fin de encontrar resultados sobre implementaciones de nuevos comandos de voz y qué tecnologías se emplearon en esos casos.
- **Diseño y construcción:** En esta fase se creó un árbol conversacional capaz de prever todos los puntos de contacto entre el usuario y el asistente inteligente, así como las respuestas a preguntas mal formuladas o incluso a insultos, para minimizar los posibles errores durante el proceso de conversación.
- **Implementación y pruebas:** En este paso se realizó la implementación y prueba de los nuevos comandos de voz utilizando las siguientes tecnologías de Google: un simulador capaz de emular un entorno de prueba deseado, un dispositivo Google Home mini y un Smartphone.
- **Despliegue:** Corresponde a la publicación de los nuevos comandos de voz en la plataforma de Google Assistant, tomando como referencias los hitos publicados para desarrolladores de la comunidad de Google.

#### 5 Arquitectura de Desarrollo

A continuación, se presenta la arquitectura que dará solución al desarrollo de los nuevos comandos de voz para el asistente virtual de Google como se muestra en la figura 2; es importante destacar que dicha arquitectura se construyó basándose en la misma que maneja Google, por lo que es necesario atender las actualizaciones que la empresa realice en las tecnologías usadas para el desarrollo del proyecto.

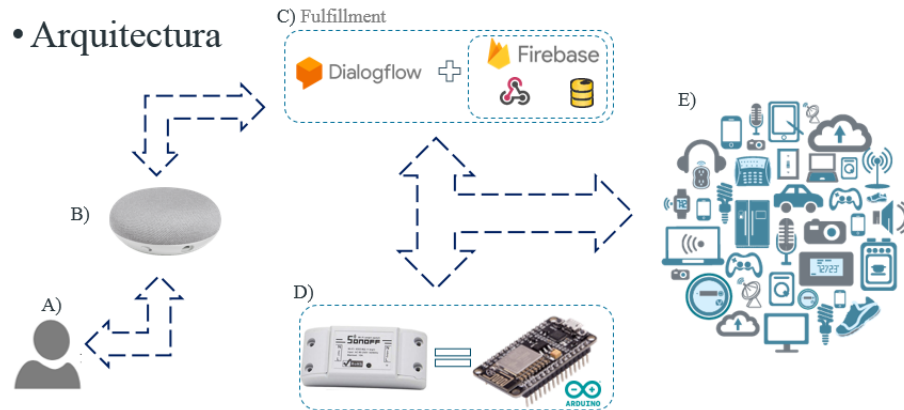


Figura 2. Arquitectura para el desarrollo de los nuevos comandos de voz.

#### A. Usuario.

Corresponde a quien utilizará el lenguaje natural para comunicarse con el dispositivo Google Home Mini, es importante resaltar que es necesario usar la frase “Ok Google” antes de alguna instrucción, para que el dispositivo reconozca cualquier cosa después de dicha oración como una orden.

La frase “OK Google”, ya se encuentra definida y es un punto clave para desencadenar la acción deseada; esta frase no se puede cambiar y es de uso exclusivo del Google Assistant.

#### B. Google Home Mini.

El dispositivo obtendrá la orden expresada por el usuario, y se encontrará conectado a la plataforma de Dialogflow. El Google Home Mini es el dispositivo físico por medio del cual se puede hacer uso del Google Assistant, dado que cuenta con hardware particular que permite una mejor captación de sonido en lugares más amplios, evitando el desfase de error por eco al momento de recibir la orden.

#### C. Fulfillment.

Es el servicio encargado de procesar el lenguaje natural de la instrucción que manda el usuario, su objetivo es encontrar en dicha instrucción, las palabras claves que sirven para identificar el comando de voz que se requiere, su acción puede ser complementada con WebHooks, como se observó en [8] y [12].

El Fulfillment es donde se manejará la mayor parte de la lógica para lograr que el comando de voz pueda hacer la interacción y/o nuevas funcionalidades que se requieran para que este pueda lograr su objetivo.

La interacción con el dispositivo de Google se llevará a cabo mediante uno o varios JSON dependiendo de los comandos y/o acciones que se requieran. El JSON es un formato de texto sencillo el cual permite el intercambio de datos de manera más accesible logrando agilizar el proceso de comunicación.



En este JSON se mostrarán las acciones que ejecutará el comando de voz; además, el archivo JSON se apoyará del lenguaje de programación que admita gRPC para especificar funcionalidades extras y tener un buen funcionamiento. En esta sección se encuentra la mayor responsabilidad sobre la arquitectura.

#### D. Sonoff y/o NodeMCU ESP8266

Es un dispositivo que actuará de intermediario para la conexión con los dispositivos externos; se planea controlar mediante comandos de voz.

Este dispositivo tiene la particularidad de permitir el entendimiento con el asistente virtual de Google y ofrecer un mejor control sobre los dispositivos externos como el Raspberry Pi [10].

Obtendrá todos los datos necesarios desde la plataforma Firebase y serán programados en el IDE de Arduino y se cargará dicha configuración al intermediario.

#### E. Dispositivos externos.

Estos realizarán una acción dependiendo del comando de voz que se haya ejecutado, por ejemplo, sea el caso de que este dispositivo este representado por un Smartphone, una acción válida para tal, será encenderse o apagarse.

## 6 Desarrollo Del Comando De Voz

La programación se realiza utilizando la plataforma Firebase mediante el archivo de "index.js", esto permite almacenar las variables en donde los dispositivos externos se enlazan para cambiar el estado de estas.

Para llevar a cabo la ejecución del comando de voz de una manera satisfactoria, se enlaza el archivo "index.js" con la plataforma Dialogflow (Fulfillment) para efectuar las funcionalidades y/o actualizaciones indicadas por el comando de voz.

Las variables que se estarán almacenando en la base de datos de Firebase en tiempo real se visualizan en la figura 3.

```
control-seminario
├── accionalarma: "apagar"
├── accioncocina: "apagar"
└── accionsala: "apagar"
```

Figura 3. Variables almacenadas en Firebase para dispositivos externos.

El comando de voz envía y actualiza las variables de estado en la plataforma de Firebase, es necesario instalar el IDE de Arduino, ya que este nos permitirá programar de manera directa la placa NodeMCU ESP8266, representada en la figura 4.

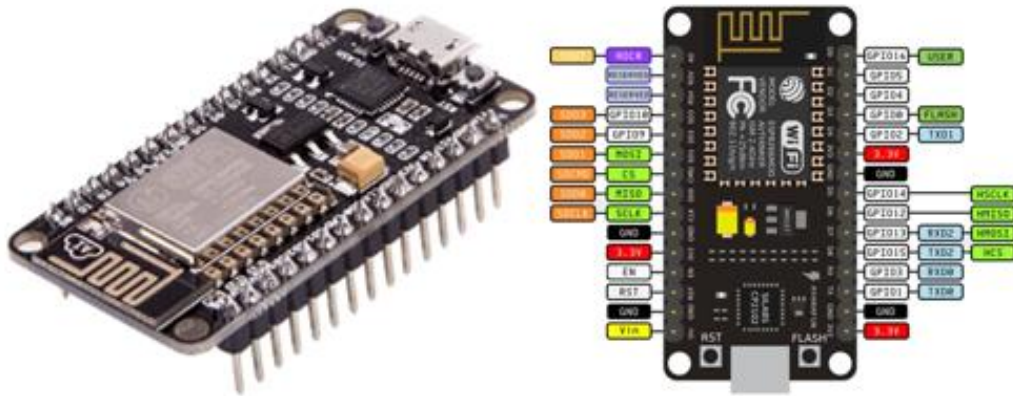


Figura 4 Placa NodeMCU ESP8266 [13]

A continuación, se muestra el esquema de conexión del control de los dispositivos externos en la Figura 5.

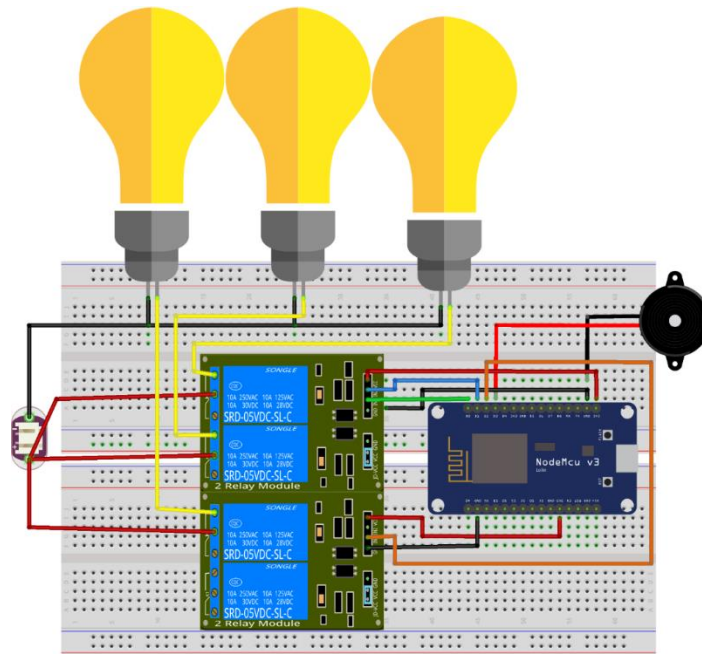


Figura 5 Esquema de conexión.

Hecha la conexión de los circuitos y componentes, se programa la placa NodeMCU ESP8266. Dentro del archivo se deben usar las bibliotecas que necesita Firebase para tener comunicación con el IDE de Arduino. Se debe crear una conexión con una red que proporcione Internet, a su vez conectar en el proceso con la base de datos de Firebase la cual cuenta con un dominio, así como una contraseña de protección para evitar filtraciones de información.

Una vez hechas las conexiones necesarias, se debe obtener las variables almacenadas de la figura 5, para que la lógica de encendido y apagado se lleve a cabo de forma adecuada.

La placa NodeMCU ESP8266 estará recuperando los valores almacenados en la base de datos de Firebase en tiempo real, esto provocará que cada vez que haya un cambio provocado por la

interacción entre el usuario y el asistente virtual de Google usando el comando de voz se efectuó al instante sin ninguna demora.

La figura 6 muestra los pasos en que se ejecuta un comando de voz para lograr un funcionamiento correcto.

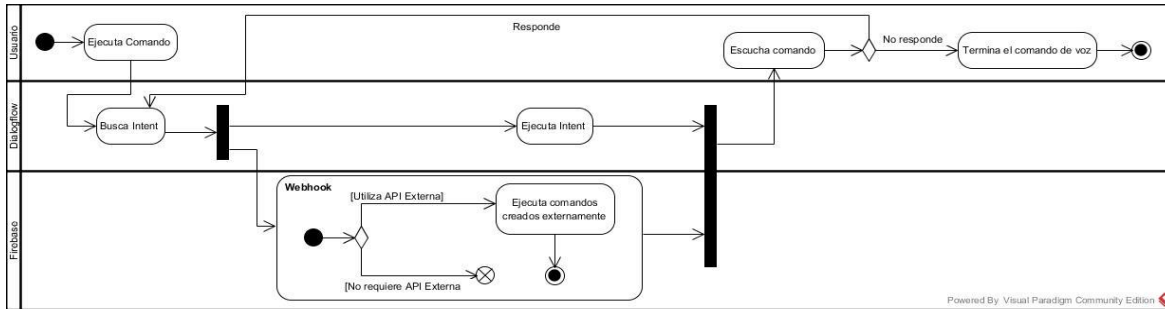


Figura 6 Funcionamiento del comando de voz.

El usuario mediante el habla inicia el nuevo comando de voz, el asistente virtual de Google buscará en la plataforma de Dialogflow el intent que corresponda, dependiendo del intent que encuentre utilizará o no la plataforma de Firebase, en Firebase se hará uso de una API externa o no, se ejecutará el intent correspondiente y, una vez ejecutado el intent, el asistente virtual de Google enviará una respuesta adecuada, donde el usuario podrá proseguir a la respuesta del asistente virtual o quedarse callado. Si el usuario contesta se volverá a buscar otro intent que se adecue a lo dicho y ejecutará el flujo de nuevo, si el usuario no respondió terminará el comando de voz.

## 7 Resultados

Mayormente, el uso de dispositivos externos se encuentra limitado por solo una acción, esto quiere decir que, cuando el usuario invoca un comando mediante “OK Google” seguido de la acción que desea, esta se lleva a cabo únicamente en un solo dispositivo, por lo que en este caso de estudio se logró superar esta limitante debido a que un comando de voz se ejecuta en más de un dispositivo externo. La Figura. 5 nos permite entender el funcionamiento y el control del NodeMCU ESP8266 sobre cada uno de los dispositivos externos (Foco de la sala, Foco de la cocina, Foco de la alarma, Zumbador de la alarma). Se optó usar el esquema de la Figura 5 debido a que así se tiene un mejor control sobre cada dispositivo, además de que es posible contralar cada acción sobre este y así se evitan errores al momento de ejecutar acciones. Para iniciar el comando de voz es necesario que el usuario pronuncie la palabra clave “OK Google” para desencadenar la acción del comando. Una vez que el comando de voz es reconocido, el usuario tendrá total control de los dispositivos disponibles. Suponiendo que el usuario enciende la luz de la sala, el dispositivo con el asistente virtual de Google responderá de manera visual y/o mediante voz para permitir una interacción clara y concisa con el usuario.

El comando de voz no se dejará de ejecutar; esto permite al usuario poder encender otro dispositivo que se encuentre a su disposición, proporcionando el control de más de un dispositivo y otorgando una gran flexibilidad en el manejo de estos. A continuación, en las figura 7 y 8 se los comandos de “encender y apagar” que demuestran el funcionamiento en cualquier dispositivo que cuente con el asistente virtual de Google.

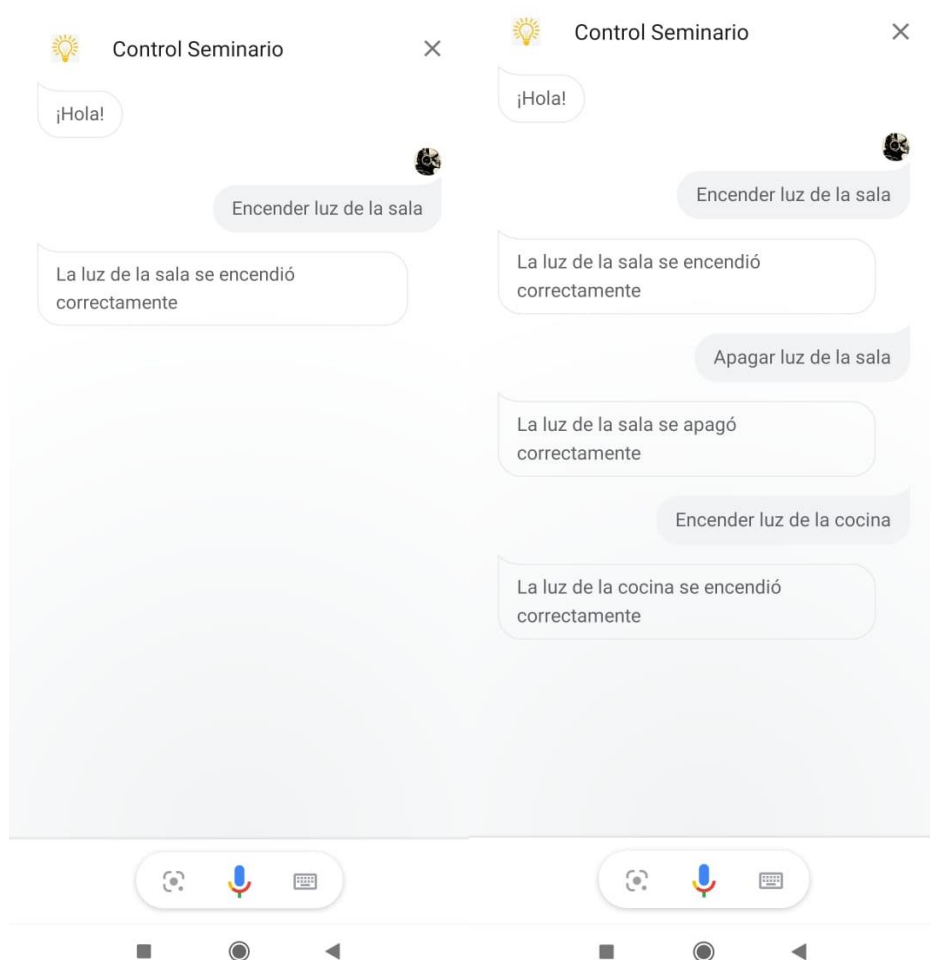


Figura 7 Funcionamiento del comando de voz en dispositivo Móvil parte 1.

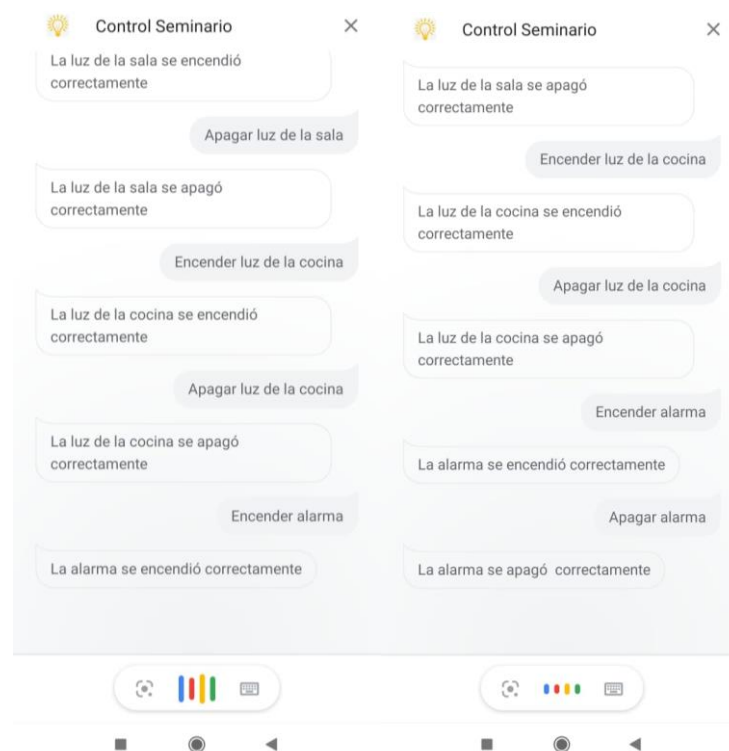


Figura 8 Funcionamiento del comando de voz en dispositivo Móvil parte 2.

Siguiendo con la demostración del uso de los nuevos comandos de voz, se presentarán las imágenes que demuestran el funcionamiento de los dispositivos externos y la forma en que estos interactúan como se muestra en la figura 9.



Figura 9 Prueba de funcionamiento de dispositivos.

Una vez terminado el desarrollo e implementación de nuevos comandos de voz, se hizo un estudio estadístico para medir la experiencia de usuario sobre los estos. Dicho análisis se realizó tomando una muestra de 20 personas, donde diez son expertos en el campo del uso de asistentes virtuales y el resto no lo son.

A continuación en la figura 10, se muestra la gráfica que representa la ponderación con respecto a la experiencia de usuario, donde diez es el grado máximo de gusto y aceptación. En dicha grafica es posible apreciar que la media exponencial es de 7.5, lo que significa que la mayoría de los usuarios tuvieron una experiencia positiva al utilizar los nuevos comandos de voz.



Figura 10 Prueba de funcionamiento de dispositivos.

## 8 Conclusiones

El desarrollo e implementación de un conjunto de instrucciones para el asistente virtual Google Assistant, es un tópico que está creciendo a pasos agigantados actualmente y lo proveen muy pocas plataformas. El utilizar las tecnologías Firebase, el agente inteligente de Dialogflow, en conjunto como la programación en Arduino de la placa NodeMCU ESP8266 permitió al usuario final un mejor control de los dispositivos externos a la hora de interactuar con éstos.

La revisión de los trabajos relacionados permitió identificar que el desarrollo e implementación de un conjunto de instrucciones en los asistentes virtuales es un tema poco abordado, debido a que sólo se analizan algunas tecnologías involucradas en la interacción con estos. Con esta investigación se beneficiarán los desarrolladores que aborden la necesidad de crear e implementar nuevos comandos de voz, ya que podrán observar las ventajas y desventajas de cada tecnología utilizada, por ejemplo, en el caso del ESP8266, el desarrollador siempre debe tener en cuenta que el tiempo de espera entre la recepción de datos y la generación de una respuesta varía según la conexión a internet, esto puede representar una desventaja para casos donde la interacción con el asistente se requiera de manera inmediata.

Debido a que la estructura de la arquitectura, así como las tecnologías que usa Google Assistant y a su vez la forma en que desarrolla los comandos de voz el asistente virtual de Google va evolucionando conforme pasa el tiempo, abriendo una gran posibilidad de crear nuevas funcionalidades que hasta el día de hoy no se puedan realizar.

## Agradecimientos

Los autores agradecen el financiamiento del Consejo Nacional de Ciencia y Tecnología (CONACyT) y el apoyo del Tecnológico Nacional de México.

## Referencias

- [1] A. Campos Albuixech, “Asistente virtual en Telegram para acceder a la información económica municipal del Ajuntament de València”, oct. 2018.
- [2] S. Junestrand, X. Passaret, y D. Vázquez, Domótica y hogar digital. Editorial Paraninfo, 2004.
- [3] E. Méndez y J. A. Moreiro González, “Lenguaje natural e indización automatizada”, Ciencias de la Información, 1999. [En línea]. Disponible en: <http://eprints.rclis.org/12685/>. [Consultado: 07-abr-2019].
- [4] M. B. Hernández y J. M. Gómez, “Aplicaciones de Procesamiento de Lenguaje Natural”, Rev. Politécnica, vol. 32, núm. 0, jul. 2013.
- [5] P. Lara Navarra y J. A. Martínez Usero, Agentes inteligentes en la búsqueda y recuperación de información. Planeta UOC, 2004.
- [6] Google, “Actions on Google Glossary | Actions on Google”, Google Developers. [En línea]. Disponible en: <https://developers.google.com/actions/glossary>. [Consultado: 15-mar-2019].
- [7] Google, “Google Assistant SDK”, Google Developers. [En línea]. Disponible en: <https://developers.google.com/assistant/sdk/overview>. [Consultado: 14-mar-2019].
- [8] G. Iannizzotto, L. L. Bello, A. Nucita, y G. M. Grasso, “A Vision and Speech Enabled, Customizable, Virtual Assistant for Smart Environments”, en 2018 11th International Conference on Human System Interaction (HSI), 2018, pp. 50–56.
- [9] A. H. Michaely, X. Zhang, G. Simko, C. Parada, y P. Aleksic, “Keyword spotting for Google assistant using contextual speech recognition”, en 2017 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), 2017, pp. 272–278.
- [10] C. Peng y R. Chen, “Voice recognition by Google Home and Raspberry Pi for smart socket control”, en 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI), 2018, pp. 324–329.
- [11] I. Hwang, J. Jung, J. Kim, Y. Shin, y J. Seol, “Architecture for Automatic Generation of User Interaction Guides with Intelligent Assistant”, en 2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA), 2017, pp. 352–355.
- [12] A. Cheng, V. Raghavaraju, J. Kanugo, Y. P. Handrianto, y Y. Shang, “Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes”, en 2018 15th IEEE Annual Consumer Communications Networking Conference (CCNC), 2018, pp. 1–5.
- [13] «Nodemcu V3 Lolin - Wifi Esp8266/esp12e - Tarjeta Desarrollo - \$ 5.402». [https://articulo.mercadolibre.cl/MLC-452860548-nodemcu-v3-lolin-wifi-esp8266esp12e-tarjeta-desarrollo-\\_JM](https://articulo.mercadolibre.cl/MLC-452860548-nodemcu-v3-lolin-wifi-esp8266esp12e-tarjeta-desarrollo-_JM) (accedido sep. 09, 2020).