

CIATEQ, A. C. Centro de Tecnología Avanzada
Gerencia de Posgrado



*La prealimentación como técnica de diseño de experiencia
de usuario aplicada a indicadores vitales propios de un
panel de instrumentos automotriz manejado por asistente
virtual*

TESIS QUE PRESENTA

Ing. Ricardo Hernández Mejía
Asesor: Dr. Francisco Javier Ibarra Villegas

Para obtener el grado de

Maestro en
Sistemas Inteligentes Multimedia

Guadalajara, Jalisco
diciembre, 2023

CARTA DE LIBERACIÓN DEL ASESOR



**GOBIERNO DE
MÉXICO**



CONAHCYT
CONSEJO NACIONAL DE HUMANIDADES
CIENCIAS Y TECNOLOGÍAS



CIATEQ

Guadalajara, Jalisco, 11 de Octubre de 2023

Mtro. Geovany González Carlos
Gerencia de Posgrado
CIATEQ, A.C.

El que suscribe, miembro del Comité Tutorial del ingeniero Ricardo Hernández Mejía, una vez revisado su Proyecto Terminal de tesis, titulado "La prealimentación como técnica de diseño de experiencia de usuario aplicada a indicadores vitales propios de un panel de instrumentos automotriz manejado por asistente virtual" autorizo que el citado trabajo sea presentado por el alumno para su revisión, con el fin de alcanzar el grado de maestro.

Sin otro particular por el momento, agradezco la atención prestada.

Atentamente

Dr. Francisco Javier Ibarra Villegas
Asesor Académico

F31e Revisión: 01-Jun-2021



CARTA DE LIBERACIÓN DEL REVISOR



**GOBIERNO DE
MÉXICO**



CONAHCYT
CONSEJO NACIONAL DE HUMANIDADES
CIENCIAS Y TECNOLOGÍAS



CIATEQ

Ciudad, Estado, 4 de diciembre del 2023.

Mtro. Geovany González Carlos
Gerencia de Posgrado
CIATEQ, A.C.

Por medio de la presente me dirijo a usted en calidad de Revisor del proyecto terminal del (la) alumno (a) **Ricardo Hernández Mejía**, cuyo título es:

"La prealimentación como técnica de diseño de experiencia de usuario aplicada a indicadores vitales propios de un panel de instrumentos automotriz manejado por asistente virtual"

Después de haberlo leído, corregido e intercambiado información con el (la) alumno(a), y realizado los cambios que le fueron sugeridos, puede ser autorizada su impresión, a fin de que se inicien los trámites correspondientes para su defensa.

Sin otro particular por el momento, y en espera de que mis sugerencias sean tomadas en cuenta en beneficio del estudiante y la Institución, agradezco la atención prestada.

Atentamente,

Mtro. Pedro Leopoldo Martínez Quintal

Grado, nombre completo

F31b Revisión: 04-Feb-2022

DEDICATORIA

Le dedico este trabajo a mi hijo Ricardo Hernández Garduño. A pesar del distanciamiento físico y la intermitente comunicación todo este tiempo, consecuencias de las decisiones de tomadas por adultos en su camino accidentado por la vida, jamás dejó de ser el motor de mi vida y la razón por la que hiciera las cosas.

Este trabajo es una muestra fehaciente de la resiliencia con la que contamos las personas ante situaciones complicadas de la vida, buscamos respuestas a incógnitas, soluciones a problemas, camino que estamos destinados a recorrer quienes poseen el valor y la curiosidad de saber: "Nosce te Ipsum", "Sapere aude", "Carpe Díem", y "Cuando llegues a la cima, sigue subiendo".

AGRADECIMIENTOS

A mi asesor de tesis, Doctor Francisco Javier Ibarra Villegas, por el tiempo dedicado a orientar, dar forma y dirección al trabajo plasmado aquí, así como al trabajo relacionado con la tesis presentado en el congreso CICITEC 2023. Los estándares de profesionalismo y calidad logrados en ambos trabajos son gracias a él.

A todos los docentes del programa de maestría CIATEQ por haber formado parte fundamental en la maduración del conocimiento y su aplicación práctica para la solución de problemas en la industria.

A las licenciadas Gabriela Soto y Sandra Rascón por haber dado seguimiento puntual a todas las necesidades y problemáticas administrativas a las que nos enfrentamos los miembros de mi generación.

A todos los conocidos, amistades y familiares que consideraron plausible y brindaron apoyo a este trabajo, a través de su valiosa atención, escucha, comentarios y aportaciones.

A Continental Automotive Occidente y CIATEQ A.C. por la vinculación y el convenio vigentes, el apoyo económico y administrativo brindado para poder realizar los estudios correspondientes al posgrado en Sistemas Inteligentes Multimedia y llevarlos a buen término.

* * *

RESUMEN

Este trabajo se originó del creciente interés por parte de diferentes industrias para implementar soluciones de asistente virtual basado en voz impulsadas por el campo de estudio del Procesamiento del Lenguaje Natural. Este trabajo está enfocado en los productos relacionados a la Interfaz Humano Máquina de la industria automotriz, específicamente el Panel de Instrumentos. Hoy en día las personas usan constantemente asistentes virtuales como Google Assistant, Alexa, Cortana o Siri en sus dispositivos electrónicos. Más aún, 31% de los autos tienen un asistente virtual integrado, por ejemplo, Ford usa Alexa, Mercedes-Benz y Hyundai usan Google Assistant, BMW y Nissan usan Cortana, GM usa IBM Watson, Honda usa Hana y Toyota usa YUI. Aparte de las soluciones de marca registrada descritas anteriormente, también hay soluciones genéricas de código abierto contemporáneas disponibles en el mercado, tales como Mycroft AI que se hace notar por sobre otras tecnologías por características como listo para usar, bien documentada, instalación simple en una PC Linux o RPI SoC, y una ejecución simple.

Este trabajo presenta una manera de usar Mycroft AI como una alternativa para agregar asistencia de voz basada en inteligencia artificial a aplicaciones en el dominio automotriz. El módulo de comunicación de voz presentado aquí maneja información relacionada a tres diferentes entidades: cinturón de seguridad, nivel de gasolina y nivel de batería, todos ellos son indicadores virtuales presentes en cualquier Panel de Instrumentos Automotriz. Dado que el enfoque de diseño de Mycroft AI se basa en Diseño Centrado en el Human (HCD), el módulo de comunicación por voz presentado aquí provee un diseño basado en experiencia de usuario (UX) real con características de prealimentación y retroalimentación.

Como conclusión, Mycroft AI demuestra gran potencial como una alternativa para agregar asistencia de voz a los productos relacionados a Interfaz Humano Máquina de la industria automotriz. Acerca del trabajo a futuro, debido al hecho que Mycroft AI está basado en Python, existen vastas posibilidades para conectar y expandir el módulo de comunicación por voz a través del uso de innumerables bibliotecas de Python para importar y procesar cualquier tipo de información, en cualquier

formato o fuente, por ejemplo la información proveniente de tecnologías de comunicación tales como CAN, LIN, ETH, MOST, GPS, V2X o cualquier otro dispositivo o tecnología para crear soluciones automotrices integrales.

Palabras clave: Panel de instrumentos; Asistente virtual; Módulo de comunicación por voz; Mycroft AI; Diseño centrado en el humano; Experiencia de usuario.

ABSTRACT

This work was originated from the increasing interest in several industries to implement voice based virtual assistant solutions powered by the Natural Language Processing field of study. This work is focused on the automotive industry Human Machine Interface related products, specifically the Instrument Panel. Nowadays people are constantly using virtual assistants like Google Assistant, Alexa, Cortana or Siri on their electronic devices. Furthermore, 31% of cars have a built-in virtual assistant, for example Ford uses Alexa, Mercedes-Benz and Hyundai use Google Assistant, BMW and Nissan use Cortana, GM uses IBM Watson, Honda uses Hana and Toyota uses YUI. Apart from the proprietary solutions described earlier, there are also contemporary open-source generic solutions available on the market, such as Mycroft AI which stands out from other technologies due to ready to deploy, well documented, simple installation on a Linux PC or RPI SoC, and simple execution.

This work presents a way to use Mycroft AI as an alternative to add artificial intelligence-based voice assistance to applications in the automotive domain. The voice communication module presented here drives information related to three different entities: seat belt, fuel level and battery level, all of them are telltales present in any automotive Instrument Panel. Since the Mycroft AI design approach is based on Human Centered Design (HCD), the voice communication module presented here provides real user experience (UX) based design showing feedforward as well as feedback features. As a conclusion, Mycroft AI demonstrates great potential as an alternative to add artificial intelligence based voice assistance to automotive industry Human Machine Interface related products. About future work, due to the fact that Mycroft AI is based on Python, there are vast possibilities for connecting and expanding the voice communication module by using countless Python libraries in order to import and process any type of information, in any format or source, for example the information from communication technologies like CAN, LIN, ETH, MOST, GPS, V2X or any other device or technology in order to create comprehensive automotive solutions.

Keywords: Instrument panel; Virtual assistant; Voice communication module; Mycroft AI; Human centered design; User experience.

ÍNDICE DE CONTENIDO

RESUMEN	v
ABSTRACT	vii
ÍNDICE DE CONTENIDO	viii
ÍNDICE DE FIGURAS.....	x
ÍNDICE DE TABLAS	xii
SIGLAS.....	xiii
1. INTRODUCCIÓN	1
1.1. ANTECEDENTES	1
1.2. DEFINICIÓN DEL PROBLEMA	2
1.3. JUSTIFICACIÓN	2
1.4. OBJETIVOS.....	3
1.4.1. Objetivo general.....	3
1.4.2. Objetivos específicos	3
1.5. HIPÓTESIS	3
2. MARCO TEÓRICO	5
2.1. ENCUESTAS Y REPORTES DE INVESTIGACIÓN SOBRE ASISTENTES VIRTUALES	5
2.2. TRABAJOS ACADÉMICOS SOBRE ASISTENTES VIRTUALES	6
2.3. REPORTE DE PATENTES SOBRE ASISTENTES VIRTUALES	9
2.4. FERIA INTERNACIONAL DE ELECTRÓNICA DE CONSUMO (CES)	13
2.5. PATENTES Y PUBLICACIONES RELEVANTES SOBRE ASISTENTES VIRTUALES	13
2.6. DISEÑO BASADO EN EXPERIENCIA DE USUARIO	14
2.7. ASISTENTES VIRTUALES.....	19
3. PROCEDIMIENTO	23
3.1. SELECCIONAR INDICADORES VITALES DE UN PANEL DE INSTRUMENTOS AUTOMOTRIZ.....	23
3.2. INVESTIGAR Y SELECCIONAR ASISTENTES VIRTUALES DE FUENTE ABIERTA	25
3.2.1. ¿Por qué usar Mycroft AI?.....	29
3.2.2. Instalación de Mycroft AI en Linux PC	30
3.2.3. Ejecución de Mycroft	30
3.2.4. Emparejar Mycroft AI para Linux.....	32
3.2.5. Mycroft modular	36

3.3. CREAR LA FUNCIONALIDAD DE PREALIMENTACIÓN PARA ASISTENTE VIRTUAL	38
3.3.1. Mycroft Skill Kit	41
3.3.2. Directorio Local	43
3.3.3. Archivos intent, entity y dialog	43
3.3.4. Archivo __init__.py	47
3.3.5. Archivo README.md	51
3.3.6. Archivo settingsmeta.yaml	52
3.3.7. Archivo manifest.yml	52
3.4. CASOS DE PRUEBA.....	52
4. RESULTADOS	54
CONCLUSIONES	56
RECOMENDACIONES	57
APORTACIÓN DE LA TESIS.....	58
APORTACIÓN SOCIAL DE LA TESIS.....	59
REFERENCIAS.....	60

ÍNDICE DE FIGURAS

Figura 1 Cantidad de documentos por tipo y año sobre asistentes virtuales	6
Figura 2 Cantidad de documentos por editorial sobre asistentes virtuales	7
Figura 3 Cantidad de documentos por campo de estudio sobre asistentes virtuales	8
Figura 4 Países y Regiones más activas sobre asistentes virtuales	9
Figura 5 Cantidad de patentes por año sobre asistentes virtuales	10
Figura 6 Top-10 de empresas aplicantes de patentes sobre asistentes virtuales	11
Figura 7 Top-10 de empresas con propiedad de patentes sobre asistentes virtuales	11
Figura 8 Cantidad de patentes por jurisdicción sobre asistentes virtuales	12
Figura 9 Niveles de procesamiento cognitivo-emocional y las Etapas del Ciclo de Acción.....	19
Figura 10 Componentes de un asistente virtual actual	20
Figura 11 Componentes de un Asistente Virtual de Fuente Abierta.....	26
Figura 12 Command Line Interface de Mycroft AI.....	32
Figura 13 Ventana de emparejamiento para dispositivo de Mycroft AI.....	33
Figura 14 Nombre, Estado, Plataforma y Versión del dispositivo Mycroft AI.	34
Figura 15 Tipo de Voz y Wake Word del dispositivo Mycroft AI	34
Figura 16 Ubicación geográfica y física del dispositivo Mycroft AI	35
Figura 17 Preferencias, medidas, huso horario y fechas del dispositivo Mycroft AI	35
Figura 18 Directorio de skill Telltales	42
Figura 19 Directorio locale de skill Telltales.....	43
Figura 20 Archivo telltales.entity.....	43
Figura 21 Archivo telltales.intent	44
Figura 22 Archivo telltales.dialog	44
Figura 23 Diagrama de estados o modos de prealimentación.....	45
Figura 24 Archivo trip_notification.intent	45
Figura 25 Archivo trip_suggestion.intent	45
Figura 26 Archivo trip_action.intent	46
Figura 27 Secuencia de prealimentación en modo notificación.....	46

Figura 28 Secuencia de prealimentación en modo sugerencia..... 46
Figura 29 Secuencia de prealimentación en modo acción..... 47

ÍNDICE DE TABLAS

Tabla 1 Patentes y artículos de revista sobre asistentes virtuales con proactividad	14
Tabla 2 Golfos de Interacción de acuerdo con la psicología humana	16
Tabla 3 Siete Etapas de Acción	16
Tabla 4 Niveles de procesamiento cognitivo-emocionales	18
Tabla 5 Grupos de tareas del NLP	21
Tabla 6 La Inteligencia Artificial detrás de Alexa, Siri, Cortana y Google Assistant.	22
Tabla 7 Características y software base de los componentes de un Asistente Virtual de Fuente Abierta	26
Tabla 8 Comparativa de asistentes virtuales contemporáneos de fuente abierta	28
Tabla 9 Características de Mycroft AI	29
Tabla 10 Configuración de listener para la wake word Conti del dispositivo Mycroft AI.....	37
Tabla 11 Necesidades, casos de uso y ejemplos para los telltales de un IP manejado por asistente virtual	40
Tabla 12 Script de MSK para la skill Telltales	42
Tabla 13 Casos de Prueba de Skill Telltales	52

SIGLAS

ANN: Artificial Neural Network, Red Neuronal Artificial.

CAN: Controller Area Network.

CES: Consumer Electronics Show.

CMU: Carnegie Melon University.

CNN: Convolutional Neural Network, Red Neuronal Convolutacional.

DNN: Deep Neural Network.

ECU: Electronic Control Unit, Unidad de Control Electrónico.

ETH: Ethernet.

FMVSS: Federal Motor Vehicle Safety Standards, Estándares Federales de Seguridad de Vehículos Motorizado.

GPS: Global Positioning System.

HDC: Human-Centered Design, Diseño Centrado en el Humano.

HMI: Human Machine Interface, Interfaz Humano Máquina.

Infoentretenimiento: Sistemas HMI que brindan tanto información como entretenimiento.

IP: Instrument Panel, Panel de Instrumentos.

LIN: Local Interconnect Network.

ML: Machine Learning.

MOST: Media Oriented Systems Transport.

MSK: Mycroft Skills Kit.

NHTSA: National Highway Traffic Safety Administration, Administración Nacional de Seguridad del Tráfico en las Carreteras.

NLG: Natural Language Generation, Generación de Lenguaje Natural.

NLP: Natural Language Processing, Procesamiento de Lenguaje Natural.

NLU: Natural Language Understanding, Entendimiento de Lenguaje Natural.

OEM: Original Equipment Manufacturer. Fabricante Automotriz.

Prealimentación: Feedforward.

Retroalimentación: Feedback.

RNN: Recurrent Neural Network, Red Neuronal Recurrente.

Skill: Módulo de aplicación de Mycroft.

STT: Speech To Text, Voz a Texto.

TTS: Text To Speech, Texto a Voz.

UCD: User Centered Design ó Design Thinking, Diseño Centrado en el Usuario.

User Experience: Experiencia de Usuario.

V2X: Vehicle to Everything.

1. INTRODUCCIÓN

Este trabajo plantea la creación de un prototipo que ponga en práctica la prealimentación (feedforward) como uno de los principios fundamentales del diseño de experiencia de usuario (*UX, User eXperience*) aplicado a un grupo de indicadores vitales de un panel de instrumentos automotriz controlado por una herramienta de asistente virtual.

1.1. ANTECEDENTES

Desde el año 2013 brindo mis servicios profesionales a la empresa Continental Automotive Group dentro del grupo de investigación y desarrollo en dispositivos Interfaz Humano Máquina para diversas empresas armadores de automóviles. La evolución de estos sistemas, partiendo de componentes meramente mecánicos, hasta los actuales componentes virtuales, con la realidad aumentada proyectando información relevante al conductor sobre el parabrisas, es fascinante por el tipo de conocimientos, ciencia y tecnología que se emplea para su diseño, construcción y puesta en marcha.

Un caso de particular interés es la rama del conocimiento conocida como *UX* que fue inventada en 1993 por Donald Norman, científico cognitivo y trabajador de la empresa Apple, con el objetivo de fusionar los conceptos de interfaz humana y la usabilidad para todo tipo de productos tecnológicos. Sin embargo, no fue sino hasta finales de la primera década del siglo XXI que tomó relevancia en la industria automotriz con la asimilación de diferentes tipos de sistemas como infoentretenimiento, navegación, integración con dispositivos móviles, pantallas táctiles, comandos de voz, etc., y más aún con los autos de conducción autónoma, los cuales son un campo fértil para una amplia gama de productos con potencial para aplicar un verdadero diseño de experiencia de usuario, como es el caso de los asistentes virtuales.

La principal motivación para realizar este trabajo es combinar precisamente el diseño de experiencia de usuario, ideado por Donald Norman (Norman, 2013), en mi campo de conocimiento y experiencia en tecnología aplicada a la industria automotriz, empleando una de las tendencias en dispositivos de interacción humana, un asistente virtual.

1.2. DEFINICIÓN DEL PROBLEMA

El área de oportunidad identificada es precisamente la introducción de los asistentes virtuales dentro de un auto. Al provenir éstos de una industria diferente a la automotriz, sus requisitos y funciones deben ser adaptadas, sin embargo, el principal diferenciador es que esta adaptación se haga considerando el diseño basado en experiencia de usuario para lograr obtener un producto disfrutable y ameno que provea satisfacción y placer al conductor y/o pasajeros de un vehículo al momento de interactuar con él.

1.3. JUSTIFICACIÓN

Con base en los antecedentes, y por el hecho de que los actuales asistentes virtuales provienen de una industria de aplicación con necesidades muy diferentes a la automotriz, se busca introducir la funcionalidad de prealimentación, como técnica de diseño de experiencia de usuario, a un asistente virtual genérico para que brinde información de indicadores vitales de un panel de instrumentos automotriz.

En la búsqueda de posicionarse y/o mantenerse a la vanguardia en el mercado de productos automotrices con diseño de experiencia de usuario, la empresa Continental Automotive (de la cual soy actualmente colaborador) será la principal beneficiada, sin embargo, como consecuencia de este trabajo, también la industria en general y los usuarios de estos sistemas resultarían favorecidos.

1.4. OBJETIVOS

1.4.1. Objetivo general

Habilitar a un panel de instrumentos automotriz controlado por asistente virtual con la funcionalidad de prealimentación para obtener un producto basado en diseño de experiencia de usuario.

1.4.2. Objetivos específicos

- a. Seleccionar indicadores vitales de un panel de instrumentos automotriz.
- b. Investigar y seleccionar asistentes virtuales de fuente abierta.
- c. Crear la funcionalidad de prealimentación usando el asistente virtual seleccionado para los indicadores vitales escogidos de un panel de instrumentos automotriz.
- d. Simular las condiciones para probar la funcionalidad de prealimentación en el asistente virtual seleccionado para los indicadores vitales escogidos de un panel de instrumentos automotriz.

1.5. HIPÓTESIS

El desarrollo tecnológico actual provoca la inserción forzada y acelerada de productos en diversas industrias, éste es el caso de los asistentes virtuales integrados dentro de un automóvil con la finalidad solicitar y brindar información de este. No obstante, carentes de principios de diseño de UX y ajenos a la industria automotriz, representan un área de oportunidad para crear un producto disfrutable y ameno que provea satisfacción y confort al usuario al hacer uso de él.

La primera tarea es la de selección de indicadores vitales de un panel de instrumentos convencional contemporáneo se desarrollaría de manera natural, ameritando una investigación corta debido a que mi actividad diaria esta embebida en la industria automotriz, esta etapa demanda la justificación de la selección final de indicadores vitales.

Posteriormente, la segunda tarea que contempla los asistentes virtuales de fuente abierta, se contempla una investigación formal de diversas soluciones contemporáneas, la comparación y exposición de sus diferentes características, así como la justificación de la selección para su despliegue en el trabajo actual.

Con base en la solución del asistente virtual de fuente abierta seleccionada, se diseñará, construirá y probará la funcionalidad de prealimentación sobre los indicadores vitales seleccionados previamente y así brindar al asistente virtual de diseño basado en experiencia de usuario.

Finalmente se creará un ambiente de pruebas o test suite que simulará las condiciones a las que el sistema estaría expuesto en un escenario real.

2. MARCO TEÓRICO

2.1. ENCUESTAS Y REPORTES DE INVESTIGACIÓN SOBRE ASISTENTES VIRTUALES

En la encuesta efectuada en el año 2019 por el sitio Clutch.com (Arora, Athavale, Maggu, & Agarwal, 2020), se identificó que 27% de las personas con acceso a tecnología contaba con asistentes virtuales habilitados tales como Google Assistant, Alexa de Amazon, Cortana de Microsoft, Siri de Apple, el mismo documento muestra la adopción de los asistentes virtuales en diferentes dispositivos y contextos, el smartphone con 85% sobrepasa por mucho a bocinas inteligentes, tabletas, laptops, smart TVs, tecnología vestible, domótica, sin embargo y particularmente importante es la aparición del automóvil con un 31% de autos empleando un asistente virtual, no obstante empleado para tareas simples y no directamente relacionadas con la operación del vehículo, por ejemplo, operaciones multimedia (reproducción de música), hacer o recibir llamadas y mensajes, consultar direcciones o instrucciones en un mapa, entre otras acciones.

Por otra parte, el reporte de investigación (Majji & Baskaran, 2021) detalla las armadoras de automóviles más representativas que emplean asistentes virtuales tanto comerciales como de creación propia, por mencionar algunos, Ford con Alexa (Ford, 2023), Mercedes-Benz (Media, 2016) y Hyundai (Hyundai, s.f.) con Google Assistant, BMW (Boeriu, Horatiu, 2018) y Nissan (Nissan, 2017) con Cortana, GM (General Motors, 2016) con IBM Watson, Honda (Honda, 2017) con Hana, Toyota (Toyota, 2017) con YUI entre otros.

Sin embargo, y a pesar de que las armadoras de automóviles han logrado integraciones exitosas de los asistentes virtuales dentro de los automóviles, el reporte de investigación enfatiza que las capacidades limitadas de los asistentes virtuales introducidos de manera improvisada en el automóvil no lograron un grado de satisfacción remarcable que compitiera con el smartphone.

2.2. TRABAJOS ACADÉMICOS SOBRE ASISTENTES VIRTUALES

El interés creciente de la academia e investigación que genera el campo de los asistentes virtuales lo muestra el sitio *lens.org* de varias maneras, a continuación, algunas de ellas:

1. Cantidad de documentos por tipo y por año sobre asistentes virtuales

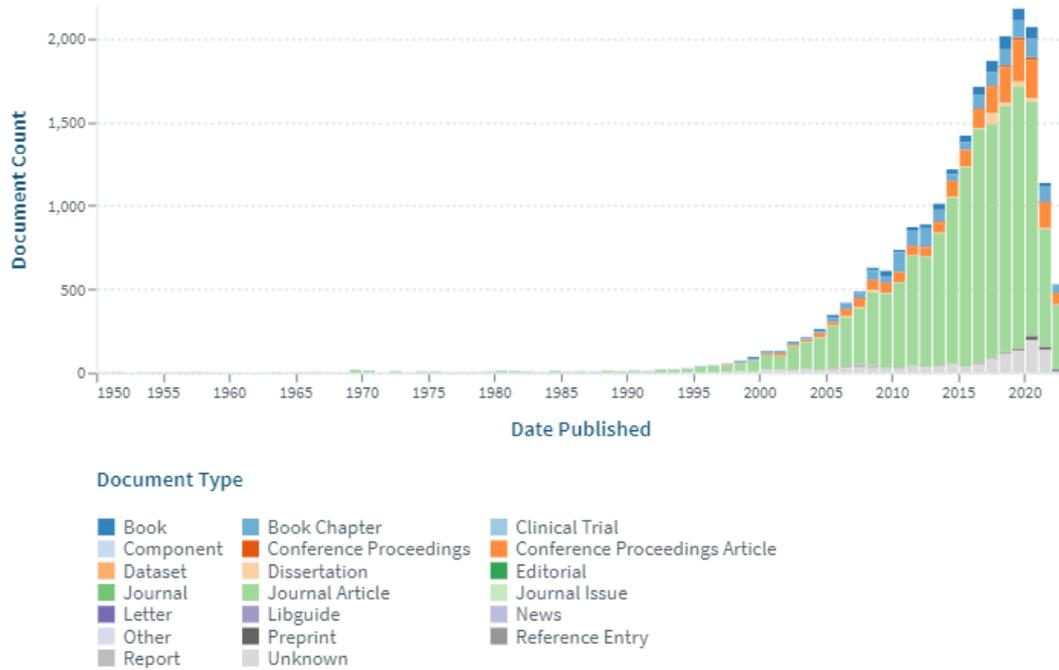


Figura 1 Cantidad de documentos por tipo y año sobre asistentes virtuales
(Lens, 2023)

La mayor cantidad de documentos académicos sobre asistentes virtuales se concentra en documentos tipo artículos de revistas, llegando a su pico a finales del año 2019, viéndose afectada en adelante por el factor externo de la pandemia por SARS-COV2.

2. Cantidad de documentos por editorial sobre asistentes virtuales

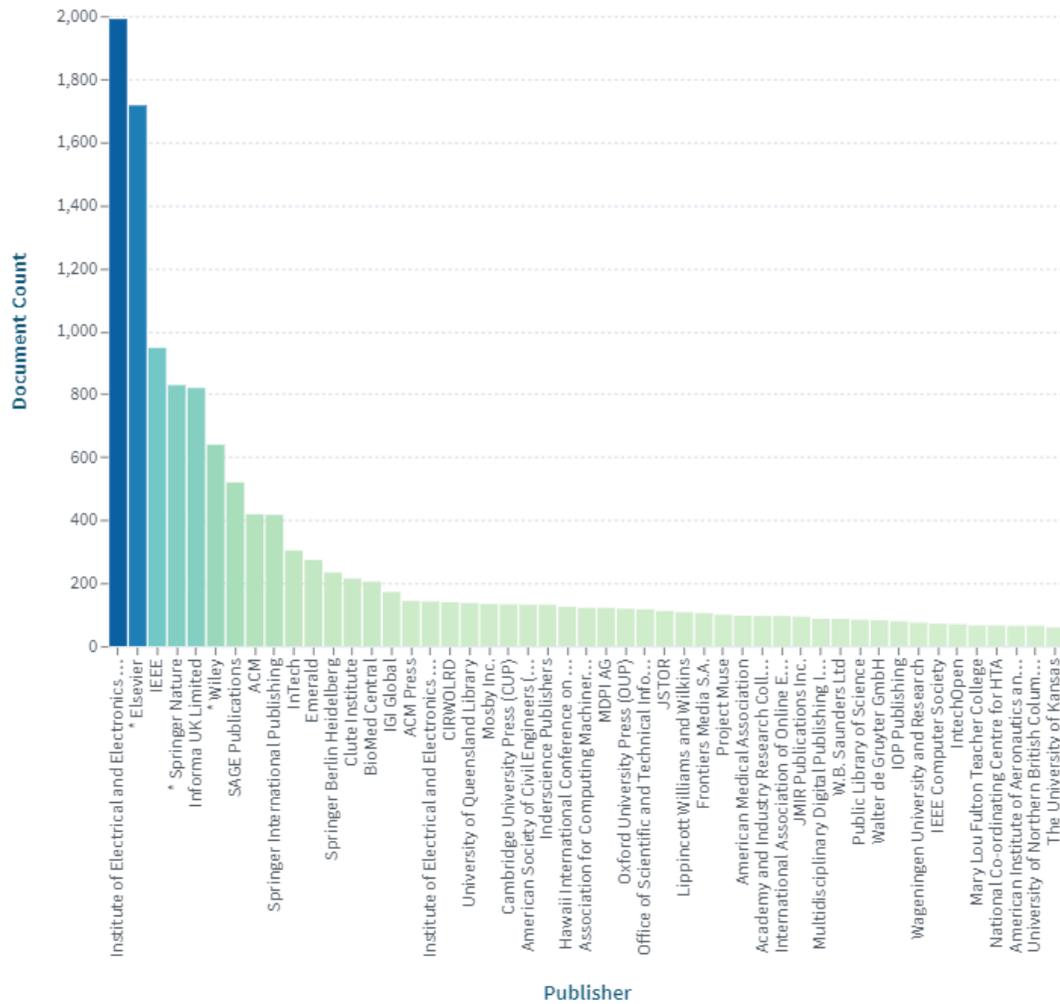


Figura 2 Cantidad de documentos por editorial sobre asistentes virtuales (Lens, 2023)

IEEE, Elsevier, Springer Nature e Informa UK Limited son las principales editoriales donde se publican documentos académicos sobre asistentes virtuales.

3. Cantidad de documentos por campo de estudio sobre asistentes virtuales

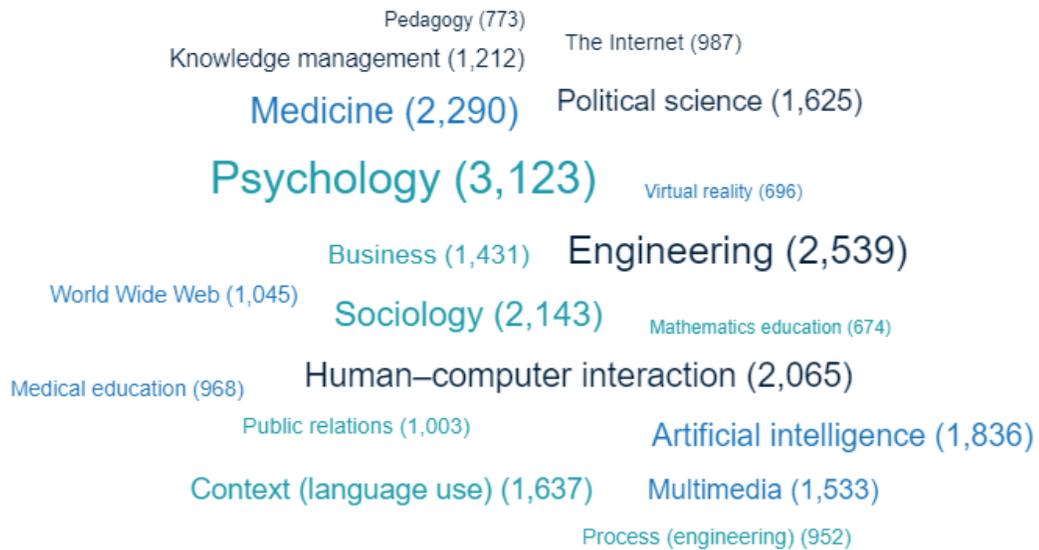


Figura 3 Cantidad de documentos por campo de estudio sobre asistentes virtuales
(Lens, 2023)

A pesar de que es esperado que campos de estudio como ingeniería, multimedia, inteligencia artificial estén relacionados con asistentes virtuales, es de resaltar el de Psicología, con más de 3000 documentos. Precisamente este campo de conocimiento relacionado a la psique humana es el factor clave que permite lograr una Interfaz Humano-Máquina o Human-Machine Interface (HMI) exitosa.

4. Países y Regiones más activas sobre asistentes virtuales

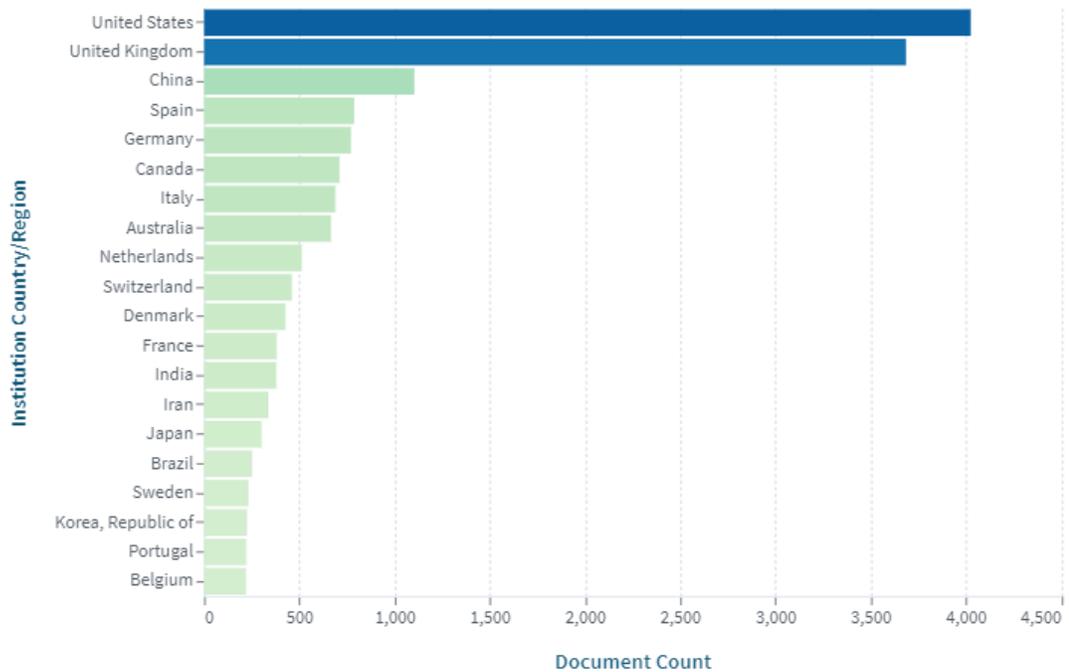


Figura 4 Países y Regiones más activas sobre asistentes virtuales

(Lens, 2023)

El par de países anglosajones Estados Unidos de América y Reino Unido se distancian mínimo por un factor de 4 en la cantidad de trabajos sobre asistentes virtuales respecto a otros países y regiones del mundo, por ejemplo, de países por excelencia y tradición automotriz como Alemania, así como de países emergentes en la industria automotriz como China.

2.3. REPORTE DE PATENTES SOBRE ASISTENTES VIRTUALES

De manera análoga a los trabajos académicos relacionados, el sitio *lens.org* muestra el interés creciente e incluso desbordado, de trabajos de patente relacionados a asistentes virtuales de varias maneras. A continuación, algunas de ellas:

1. Cantidad de patentes por año sobre asistentes virtuales

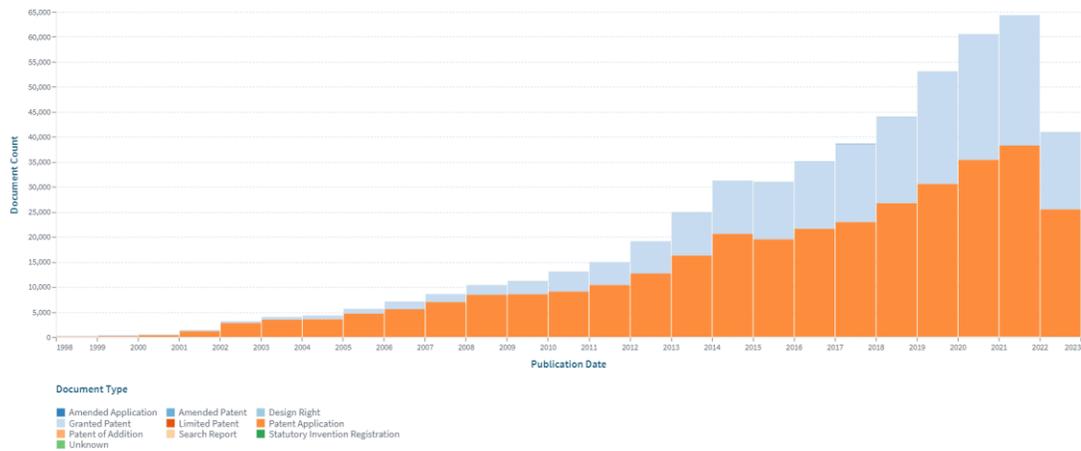


Figura 5 Cantidad de patentes por año sobre asistentes virtuales

(Lens, 2023)

La cantidad de patentes de diferentes tipos, pero especialmente las aplicaciones y las otorgadas sobre asistentes virtuales se distancia por un factor de 30 en la cantidad de trabajos académicos, naturalmente esta enorme diferencia obedece a razones con fines de lucro de las empresas tecnológicas de productos y servicios.

2. Top-10 de empresas con patentes sobre asistentes virtuales

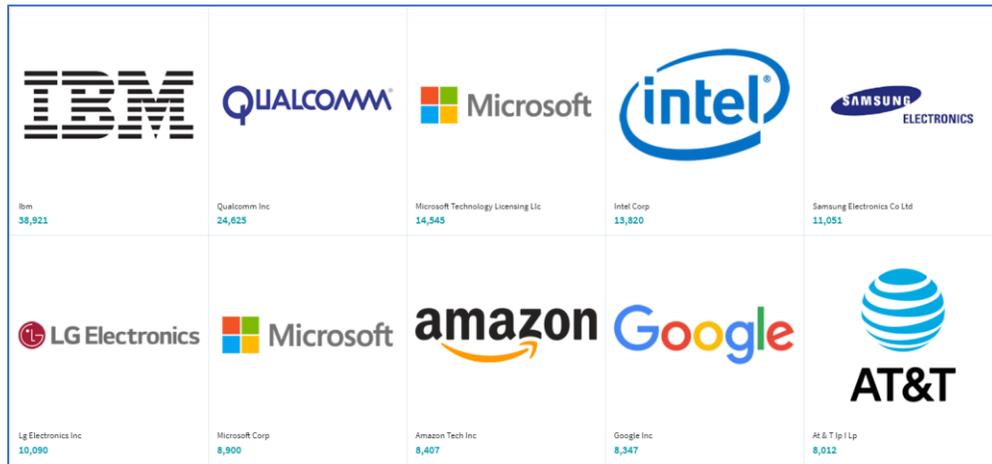


Figura 6 Top-10 de empresas aplicantes de patentes sobre asistentes virtuales (Lens, 2023)

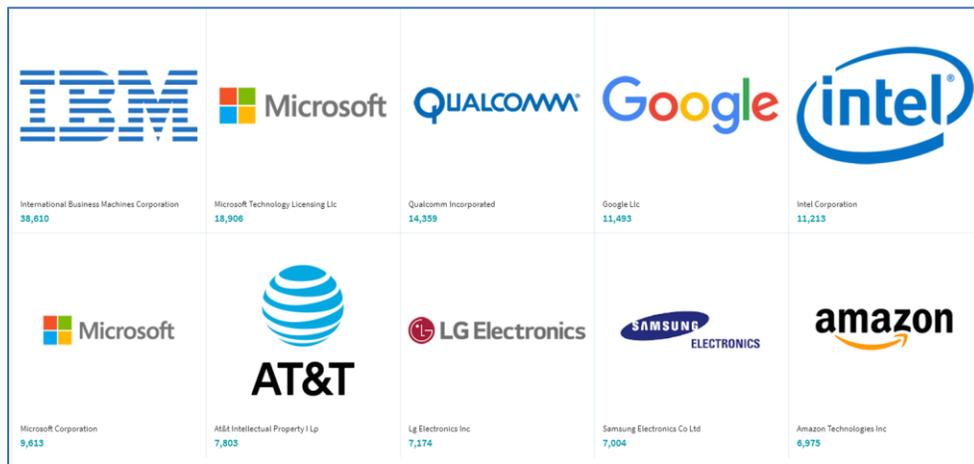


Figura 7 Top-10 de empresas con propiedad de patentes sobre asistentes virtuales (Lens, 2023)

No sorprende la presencia de empresas como Google, Amazon y Microsoft debido a sus ya bien conocidos productos: Google Assistant, Alexa y Cortana respectivamente; sin embargo, empresas relacionadas con la electrónica de consumo también tienen una presencia importante como Qualcomm, Samsung Electronics o LG Electronics. Finalmente, sí sorprende la presencia de empresas de servicios como AT&T y la predominantemente IBM.

3. Cantidad de patentes por jurisdicción sobre asistentes virtuales

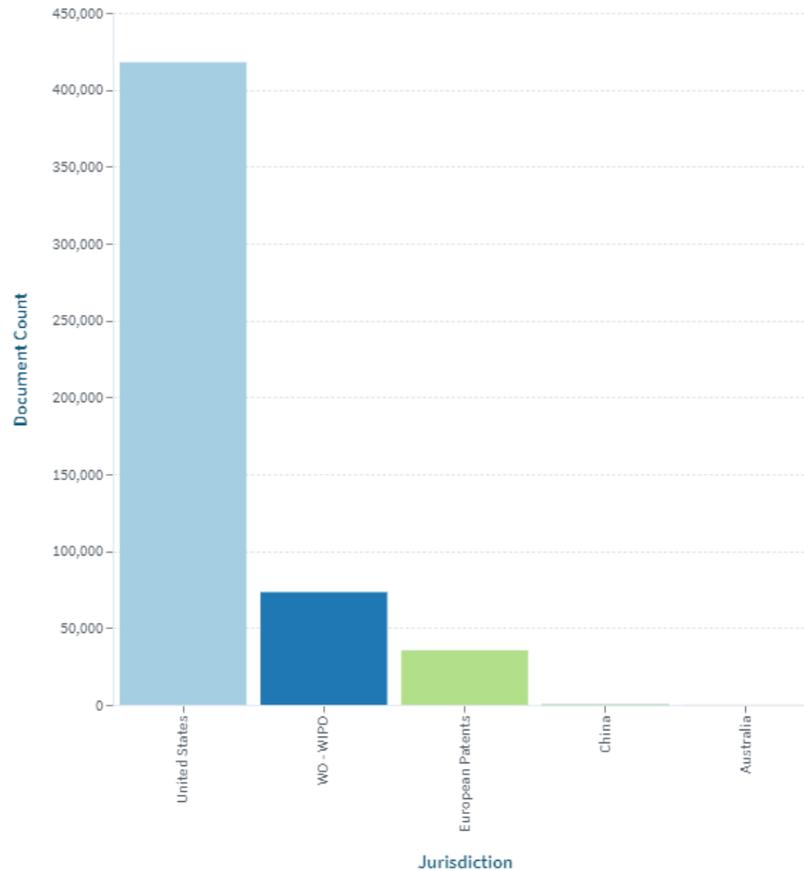


Figura 8 Cantidad de patentes por jurisdicción sobre asistentes virtuales
(Lens, 2023)

La gráfica muestra las diversas jurisdicciones o regiones de aplicación de las patentes, Estados Unidos de América es la mayor región de jurisdicción del mundo con más de 400,000 patentes, seguido por más de 70,000 patentes que aplican a nivel internacional, y en tercer lugar más de 35,000 patentes aplican para la Unión Europea, aunque en la gráfica aparecen otras regiones en realidad sus números no son significativos.

2.4. FERIA INTERNACIONAL DE ELECTRÓNICA DE CONSUMO (CES)

El evento de alcance internacional por excelencia donde cada año diversos avances y desarrollos tecnológicos se presentan por parte de las empresas más importantes del mundo pertenecientes a diversas industrias es el *Consumer Electronics Show (CES)* o Feria Internacional de Electrónica de Consumo, la cual se lleva a cabo en la ciudad de Las Vegas, en el estado de Nevada, de los Estados Unidos de América.

Precisamente en su edición 2021, la armadora Mercedes-Benz presentó su producto tecnológico automotriz más innovador hasta el momento, la Mercedes-Benz User Experience Hyperscreen (Mercedes-Benz Group Media, 2021), la cual cuenta con múltiples pantallas y un asistente virtual que además de operar comandos de voz iniciando con "Hey Mercedes", con base a su posicionamiento global, evalúa sus alrededores y brinda información relevante al conductor y/o pasajeros, tales como estaciones de servicio, restaurantes, estacionamientos, sitios de interés, entre otros, esta funcionalidad lleva por nombre: "Mercedes Travel Knowledge"; implícitamente representa la característica de proactividad agregada a un asistente virtual convencional.

2.5. PATENTES Y PUBLICACIONES RELEVANTES SOBRE ASISTENTES VIRTUALES

La característica de proactividad de los asistentes virtuales permite extender el esquema tradicional y convencional de pregunta-respuesta, la Tabla 1 muestra trabajos relevantes relacionados a esta característica.

Tabla 1 Patentes y artículos de revista sobre asistentes virtuales con proactividad

Referencia	Título	Descripción
Patente	Proactive Virtual Assistant (Estados Unidos Patente n° US 2020/0175386 A1, 2021)	Al evaluar el contenido e información del usuario, el asistente virtual pasa a una modalidad proactiva de señalización para iniciar conversaciones. Por ejemplo, un cambio de itinerario en vuelo aparte de notificar al usuario le consulta para proceder a reagendar vuelo.
Revista	The Role of Trust in Proactive Conversational Assistants (Kraus, Wagner, Callejas, & Minker, 2021)	Publica que el grado de proactividad de los asistentes virtuales tiene un impacto significativo a nivel cognitivo en la confianza del usuario en el sistema, así como efectos positivos en la experiencia de usuario, útil especialmente en situaciones para prevenir escenarios negativos. Los grados de proactividad propuestos son: Ninguno (el usuario inicia diálogos), Notificación, Sugerencia e Intervención.
Patente	Automotive Virtual Personal Assistant (Estados Unidos Patente n° US 2019/0311241 A1, 2019)	La proactividad en un asistente virtual en escenarios dependientes del contexto del usuario y del vehículo, permite predecir necesidades y acciones futuras brindando seguridad y experiencia de usuario mejorada. Por ejemplo, el asistente virtual monitorea activamente el estado del vehículo sea que esté estacionado o en movimiento para brindar notificaciones relevantes.

Elaboración propia

2.6. DISEÑO BASADO EN EXPERIENCIA DE USUARIO

El diseño basado en UX es un concepto que nació en la década de 1990 en el contexto de inicial de los sistemas con HMI, cuando Donald Norman, arquitecto de la empresa Apple dio vida al concepto de UX, en sus propias palabras: *"Inventé el término porque pensé que la interfaz humana y la usabilidad eran muy estrechas: yo quería cubrir todos los aspectos de la experiencia de una persona con un*

sistema, incluyendo diseño industrial, elementos gráficos, la interfaz, la interacción física, y la manual."

Norman fundamenta el diseño basado en UX en principios de psicología humana como la cognición, las emociones, las acciones e interacciones, principios que han permanecido y seguirán permaneciendo invariables a través del tiempo independientemente de los diferentes avances tecnológicos con los que se cuenta.

El enfoque de diseño que pone las necesidades, capacidades y comportamiento de las personas en primer lugar y sobre las necesidades, capacidades y comportamiento de los sistemas o tecnologías, es el *Diseño Centrado en el Humano* o *Human-Centered Design (HDC)*, una filosofía de diseño que al girar en torno a observar y entender a las personas, asegura la generación de productos usables y entendibles a la vez que disfrutables y deleitables que proveen placer y satisfacción al interactuar con ellos.

En el proceso de interacción entre personas y/o sistemas, la psicología humana (Norman, 2013) identifica que nos enfrentamos a dos golfos, el Golfo de Ejecución y el Golfo de Evaluación.

Tabla 2 Golfos de Interacción de acuerdo con la psicología humana

Golfo	Descripción
Golfo de Ejecución	<p>Con base a acciones se intenta conocer cómo funcionan las cosas. Es representado por el Hacer.</p> <p>Se emplean elementos:</p> <ul style="list-style-type: none"> a. <i>Significantes</i> (marcas, símbolos, indicaciones, sonidos) b. <i>Restricciones o Limitaciones</i> (rangos de operación) c. <i>Mapeos</i> (correspondencias espaciales o de forma) d. <i>Modelos Conceptuales</i> (explicación preconcebida de funcionamiento) <p>Fundamentalmente emplea la <i>Prealimentación</i>.</p>
Golfo de Evaluación	<p>Con base a resultados se intenta conocer qué sucedió. Es representado por el Interpretar.</p> <p>Fundamentalmente emplea la <i>Retroalimentación o Feedback</i>.</p>

Elaboración propia

Lo anterior da pie al modelo de interacción llamado **Las Siete Etapas de Acción** (Norman, 2013):

Tabla 3 Siete Etapas de Acción

Etapas de acción	Paso	Descripción
Etapas de Ejecución	1. Planear	La acción
	2. Especificar	Una secuencia de acciones
	3. Realizar	La secuencia de acciones
Etapas de Evaluación	4. Percibir	Qué sucedió
	5. Interpretar	Encontrar sentido a lo sucedido
	6. Comparar	El resultado contra lo esperado
	7. Meta	Objetivo a lograr de una acción

Elaboración propia

A pesar de lo simplificado del modelo de *Las Siete Etapas de Acción*, éste provee un marco de referencia útil para el entendimiento de la interacción humana, más aún, en gran medida afecta nuestro estado emocional.

Aunque tendemos a creer que los pensamientos o procesos cognitivos pueden ser separados de las emociones, esto es falso, el cerebro humano está estructurado para actuar sobre el mundo, cada acción genera expectativas y las expectativas están íntimamente relacionadas a nuestras emociones. Mientras que la cognición proporciona entendimiento de las acciones, la emoción asigna juicios de valor sobre las acciones.

Una explicación notable de la relación que guardan los procesos cognitivos y emocionales del cerebro es el modelo que considera tres niveles de procesamiento:

Tabla 4 Niveles de procesamiento cognitivo-emocionales

Nivel de procesamiento	Descripción
Visceral	<p>Nivel de procesamiento más básico, también conocido como <i>El cerebro de lagartija</i>. Permite reaccionar rápida, automática e inconscientemente, tal como un reflejo, opera a nivel de los sentidos. No requiere aprendizaje, alude al comportamiento heredado genéticamente.</p> <p>Fuertemente ligado a la musculatura del cuerpo, el sistema motor, tensión representa un estado negativo, mientras que relajado representa un estado positivo, es sinónimo de percepción inmediata, donde las apariencias, sonidos, efectos sobre la vista, tacto, olfato son lo que importa, el resultado final es la atracción o el rechazo.</p>
Comportamental	<p>Nivel de procesamiento donde residen las habilidades aprendidas impulsadas por situaciones o eventos pasados. Aunque las acciones y análisis siguen siendo inconscientes, a cada acción se le asocia una expectativa.</p> <p>La retroalimentación es un factor crítico para el manejo de expectativas, sea que se obtenga respuesta positiva, o respuesta negativa, dependiendo de la expectativa original, se provocará relajación o tensión en nosotros.</p>
Reflexivo	<p>Nivel de procesamiento de la cognición consciente en donde el entendimiento, razonamiento y toma de decisiones tienen lugar. La reacción a este nivel es lenta y profunda debido a que los procesos cognitivos que operan permiten alcanzar los niveles más altos de las emociones.</p> <p>Estados emocionales basados en causas y estimaciones de resultados conducen a por ejemplo orgullo o culpa (por una acción propia, o a elogiar o culpar a alguien (por acción de algo o alguien más).</p> <p>Los niveles altos de emociones conducen a respuestas psicológicas al límite, es así como se puede alcanzar altos grados de placer o altos grados de molestia.</p>

Elaboración propia

Norman indica que un producto con diseño basado en UX debe tener en cuenta los tres niveles de procesamiento, a esto le llama Diseño Emocional.

A manera de síntesis, la Figura 9 expone de manera clara todos los conceptos expuestos previamente y su relación con los Golfos de Interacción de acuerdo con la psicología humana.

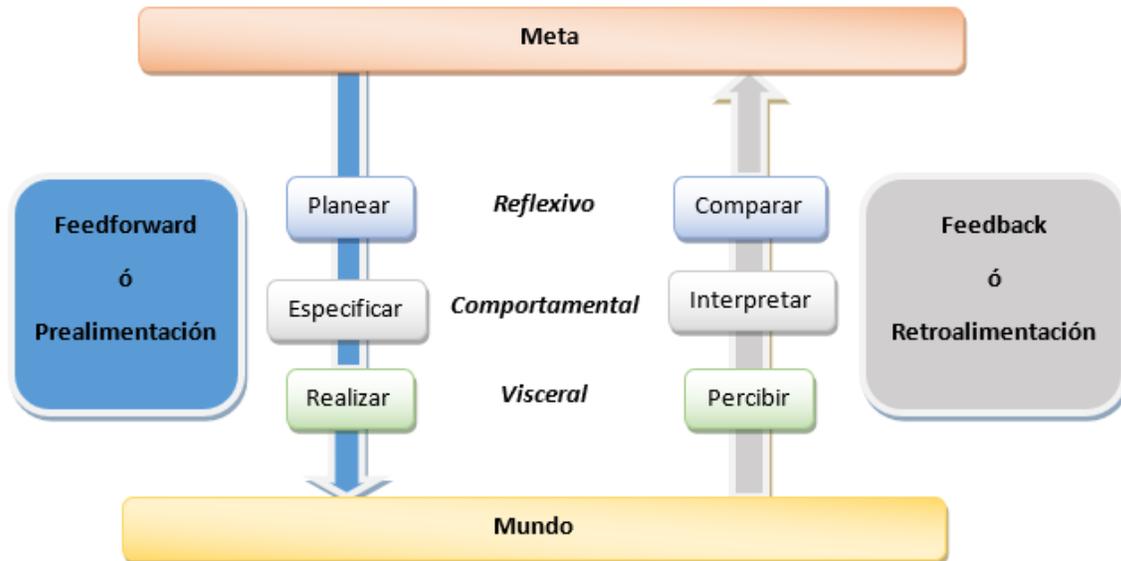


Figura 9 Niveles de procesamiento cognitivo-emocional y las Etapas del Ciclo de Acción

Elaboración propia

La información que ayuda a la ejecución es decir el *Hacer*, es la *Prealimentación*. La información que ayuda a entender qué sucedió después de la ejecución, es decir el *Interpretar*, es la *Retroalimentación*.

Cualquier producto o servicio debe contemplar los elementos antes mostrados para considerar que se empleó Diseño basado en UX en su construcción.

2.7. ASISTENTES VIRTUALES

Tecnologías de reconocimiento de voz como la ShoeBox de IBM y los sistemas de síntesis de voz como Text To Speech (TTS) nacidos en los años 60s del siglo XX, fueron los precursores de los asistentes virtuales que conocemos en la actualidad,

precisamente una arquitectura simplificada de un asistente virtual contemporáneo es el que se muestra en la Figura 10.

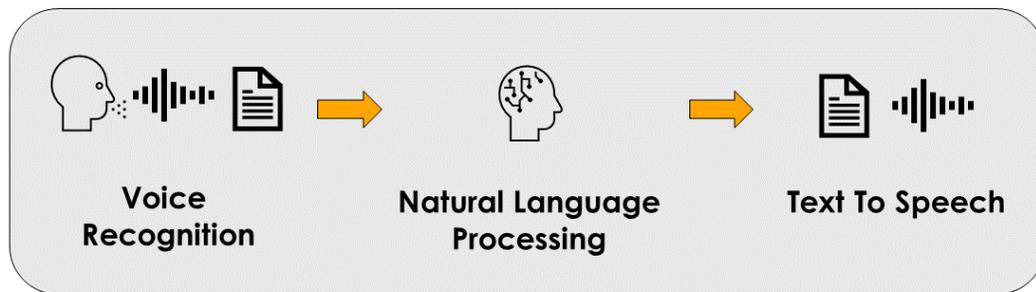


Figura 10 Componentes de un asistente virtual actual

Elaboración propia

El núcleo de los asistentes virtuales emplea el *Procesamiento de Lenguaje Natural* o *Natural Language Processing (NLP)* (Gazi Yalçın, 2021), el cual es una rama de conocimiento e investigación multidisciplinaria dividida entre ciencias de la computación, lingüística e inteligencia artificial que estudia y atiende los problemas inherentes al procesamiento y manipulación del lenguaje humano con el objetivo principal de brindar a las computadoras la habilidad o capacidad de entendimiento similar al de una persona.

Para lograr entender el lenguaje humano, el NLP hace uso de diferentes métodos basados en estadística, reglas, algoritmos y *Machine Learning (ML)* como rama de la inteligencia artificial. Debido a la complejidad sintáctica, semántica, de estructura y en ocasiones ambigua del lenguaje humano, el procesamiento se compone de diferentes grupos de tareas, descritas en la Tabla 5.

Tabla 5 Grupos de tareas del NLP

Tarea	Descripción
Morfosintáxis	Procesamiento gramatical, parseo, segmentación morfológica o por palabras, extracción de términos, descomposición de oraciones, etc.
Semántica	Lógica y significado de las palabras, Identificación y relación lógica de oraciones y temas, proveer respuestas a preguntas, Generación de Lenguaje Natural ó Natural Language Generation (NLG), análisis sentimental (emociones), desambigüedad de palabras.
Discurso	Análisis del contexto en la vida real, resumen automático de grandes entradas, resolución de coreferencias (Palabras que se refieren al mismo objeto).
Voz	Reconocimiento de voz, segmentación de voz, TTS
Diálogo	Intercambio de lenguaje tanto inicial como continuo.
Cognición	Adquisición de conocimiento y entendimiento a través del pensamiento, experiencia o sentidos. Es el nivel más abstracto usualmente es referido como Entendimiento de Lenguaje Natural ó Natural Language Understanding (NLU).

Elaboración propia

Cada grupo de tarea de la Tabla 5 representa un campo de estudio, investigación y aplicación vasto con diferentes métodos y técnicas, una de las más ampliamente usadas son las *Redes Neuronales Artificiales (ANN)* en sus derivaciones, las *Redes Neuronales Recurrentes (RNN)* y las *Redes Neuronales Convolucionales (CNN)*, ambas también se les identificadas en la literatura como *Deep Neural Network (DNN)*.

Las RNN son ideales para escenarios con datos secuenciales, datos en donde el orden en que se observan y se pretenden evaluar es importante, por ejemplo, texto, secuencias de ADN, imágenes y audio como la voz humana, entre otros.

Las CNN son uno de los modelos más poderosos de aprendizaje profundo supervisado y por esta razón se emplean en reconocimiento de imágenes, análisis de video, detecciones anómalas, procesamiento de lenguaje natural, entre otros.

La Tabla 6 de Elaboración propia muestra el resultado de investigación sobre los métodos y técnicas de inteligencia artificial empleadas en empresas puntero.

Tabla 6 La Inteligencia Artificial detrás de Alexa, Siri, Cortana y Google Assistant

Producto	Alexa	Siri	Cortana	Google Assitant
				
Empresa	<i>Amazon</i>	<i>Apple</i>	<i>Microsoft</i>	<i>Google</i>
Método/Técnica	DNN (Young-Bum, 2018)	DNN (Team, 2017)	RNN, DNN (Yong, Jinyu, Shixiong, Liping, & Yifan, 2018)	Long Short-Term Memory RNN (Beaufays, Françoise, 2015)

Elaboración propia

3. PROCEDIMIENTO

La esencia del trabajo de tesis se plasmó en el objetivo general que es: *"Habilitar un panel de instrumentos automotriz controlado por asistente virtual con la funcionalidad de prealimentación para obtener un producto con diseño basado en experiencia de usuario"*.

La secuencia de pasos fue descrita originalmente en los objetivos específicos del trabajo de tesis, los cuales son:

- a. *Seleccionar indicadores vitales de un panel de instrumentos automotriz.*
- b. *Investigar y seleccionar asistentes virtuales de fuente abierta.*
- c. *Crear la funcionalidad de prealimentación usando el asistente virtual seleccionado para los indicadores vitales escogidos de un panel de instrumentos automotriz.*
- d. *Simular las condiciones para probar la funcionalidad de prealimentación en el asistente virtual seleccionado para los indicadores vitales escogidos de un panel de instrumentos automotriz.*

A continuación, se precisarán cómo se llevaron a cabo los objetivos específicos que en conjunto permitieron alcanzar el objetivo general del trabajo de tesis.

3.1. SELECCIONAR INDICADORES VITALES DE UN PANEL DE INSTRUMENTOS AUTOMOTRIZ

Un *Panel de Instrumentos* o *Instrument Panel* (IP), en la industria automotriz también conocido como *clúster*, es en la actualidad un dispositivo computarizado conectado en red a otros dispositivos dentro del automóvil, el término técnico de los elementos de la red es *Unidad de Control Electrónico* o *Electronic Control Unit* (ECU).

La función principal del IP es la de realizar operaciones simples de manipulación de señales provenientes de mediciones hechas por diversos ECUs conectados a

sensores adquiriendo variables físicas de interés del vehículo, con la finalidad de presentar visual y auditivamente las señales de interés (Ribbens, 2017), en este trabajo se hará referencia a estas señales de interés como indicadores vitales o telltales.

Algunos de los indicadores del IP son:

- a. Nivel de combustible
- b. Nivel de batería
- c. Temperatura del motor
- d. Velocidad del vehículo
- e. Revoluciones del motor
- f. Presión de la bomba de gasolina
- g. Presión del aceite
- h. Nivel de Aceite
- i. Temperatura del anticongelante
- j. Temperatura ambiente exterior
- k. Nivel de líquido limpiaparabrisas
- l. Nivel de líquido de frenos
- m. Deslizamiento de llantas

Los Estados Unidos de América cuenta con la *Administración Nacional de Seguridad del Tráfico en las Carreteras* o *National Highway Traffic Safety Administration* (NHTSA), la cual es una agencia pública encargada de (en sus propias palabras) salvar vidas, prevenir lesiones y reducir los costos económicos debido a los choques automovilísticos, a través de la educación, la investigación, las normas de seguridad y el cumplimiento de las leyes.

Precisamente, la NHTSA es la encargada de definir los estándares de regulación de controles, señalamientos e indicadores dentro de los vehículos automotores, de lo cual se publicó en 2005 una actualización (Government, 2005), en donde se especifican requisitos relacionados a ubicación, identificación e iluminación de los elementos indicadores, dependiendo de su relevancia.

Además, con base en su anexo *Estándares Federales de Seguridad de Vehículos Motorizados* o *Federal Motor Vehicle Safety Standards (FMVSS)*, la NHTSA enlista indicadores considerados de particular relevancia para la seguridad del conductor y los pasajeros, precisando lo siguiente: “los indicadores que hayan sido activados deberán ser siempre visibles al conductor”. La lista de estos indicadores es la siguiente:

1. Fallas en el sistema de frenos
2. Falla en la bolsa de aire frontal
3. Falla en la bolsa de aire lateral
4. Presión baja de llantas
5. Apagado de bolsa de aire del pasajero
6. Luces Altas
7. Direccionales
8. Cinturón de seguridad

Para el trabajo de tesis, se consideró trabajar con un número limitado de telltales por simplicidad, y por su relevancia en pruebas estándar y de seguridad de la industria automotriz sean OEMs o proveedores como Continental, se decidió por seleccionar los siguientes telltales:

1. Cinturón de seguridad
2. Nivel de combustible
3. Nivel de batería

Estos telltales son el objetivo sobre el que el asistente virtual efectúa la función de prealimentación.

3.2. INVESTIGAR Y SELECCIONAR ASISTENTES VIRTUALES DE FUENTE ABIERTA

La arquitectura de software de un asistente virtual sea de fuente abierta o licenciada (Reid, 2020), generalmente cuenta con los componentes mostrados en la Figura 11.

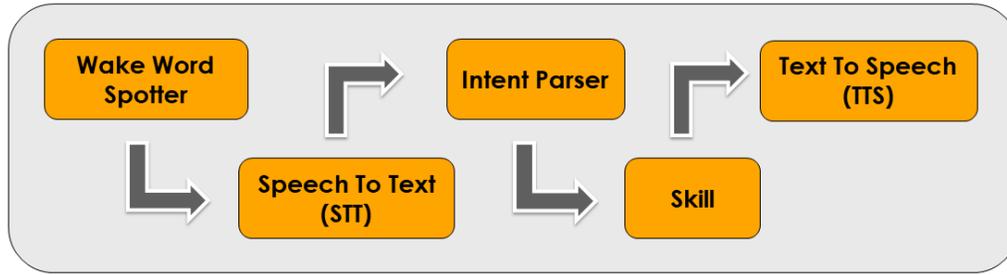


Figura 11 Componentes de un Asistente Virtual de Fuente Abierta

Elaboración propia

La Tabla 7 detalla las características de cada componente de un asistente virtual, su correspondiente descripción, así como el software base con el cual son construidos.

Tabla 7 Características y software base de los componentes de un Asistente Virtual de Fuente Abierta

Componente	Descripción	Ejemplos de Software
Wake Word Spotter	Escucha por la palabra clave (wake word o hot word), en caso de detección, activa la capa de Speech To Text (STT).	PocketSphinx, Precise, Snowboy, Porcupine.
Speech To Text (STT)	Ejecuta el reconocimiento de palabras y las traduce a texto, la salida se denomina utterance o declaración.	Kaldi, DeepSpeech, CMU Sphinx.
Intent Parser	Analiza con NLP o NLU las declaraciones del STT (tiempo, fechas u objetos) y define la acción a tomar.	Adapt, Padatious, NLTK, Rasa, Rhino.
Skill	Con base en la acción a tomar, ejecuta una skill o aplicación que crea la respuesta en formato legible por humanos con NLG.	Python, Node.js, Javascript, REST, HTTP,
Text To Speech (TTS)	La respuesta producida por el componente Skill es sintetizada por voz.	eSpeak, Festival, Mimic/Mimic2, MaryTTS

Elaboración propia

Mientras que las soluciones comerciales de asistentes virtuales demandan el pago de una licencia por uso o en su defecto la compra de sus productos con hardware específico que las mismas empresas venden, las soluciones de fuente abierta no exigen compra o pago alguno, esto es el software y hardware que emplean es genérico y de libre uso.

Para la realización de este proyecto de tesis, el uso de soluciones de fuente abierta brinda la libertad requerida para decidir entre varias opciones, por la solución que se encuentre a la mano o sea de fácil adquisición y uso, es decir cuenta con prestaciones y servicios necesarios que sean accesibles de manera clara en términos de documentación y de la manera más simple posible en términos ejecución. Características adicionales son la necesidad de conexión a internet, el enfoque a la privacidad de los datos, el nivel de personalización de la solución y qué tan populares son los correspondientes proyectos con base a su actividad en su servidor, menciones y evaluaciones positivas. La Tabla 8 generada a partir de (Reid, 2020), (Mousa, 2018), (YourTechDiet, 2022) y (Bright, Stan, 2022), muestra una comparativa entre diferentes asistentes virtuales de fuente abierta contemporáneos donde las características previamente descritas son mostradas.

Tabla 8 Comparativa de asistentes virtuales contemporáneos de fuente abierta

Asistente	OS y Hardware	Lenguaje de programación	Popularidad	Internet	Privacidad	Personalización	Documentación
Mycroft AI	Linux / RPI	Python / Bash	😊	☁️	✅	😊	😊
Leon	Windows / Linux / Mac	Node.js / Python / HTTP	😊	☁️	✅	😐	😐
Rhasspy	RPI	Docker / Python / Shell	😐	🧠	✅	😊	😞
Jasper	RPI	Python	😞	☁️	❌	😐	😞
Almond	Linux / Web	Javascript	😞	☁️	✅	😐	😞
OpenAssista nt	Windows / Linux / Mac	Software Development Kit propio	😞	🧠	✅	😐	😞
LinTO	RPI	Docker / Python / Bash / C++ / Java	😐	☁️	✅	😐	😊
Aimybox	Android / iOS	Apache 2.0	😐	☁️	❌	😐	😞
Kalliope	RPI / Linux / Android	Python / REST / Bash	😐	🧠	❌	😐	😊

😊 alta, 😐 regular, 😞 baja, 😞 pésima, ✅ sí, ❌ no, ☁️ acceso a internet, 🧠 autónomo

Elaboración propia

3.2.1. ¿Por qué usar Mycroft AI?

Mycroft AI cuenta con todos y cada uno de los componentes de software definidos en la arquitectura de un asistente virtual de fuente abierta de la Figura 11, por lo que cumple con los requisitos funcionales para el desarrollo de la tesis respecto al reconocimiento, procesamiento y síntesis de voz.

En (Mycroft AI, 2020) sus creadores enlistan en la Tabla 9 las principales razones para decidirse a usarlo:

Tabla 9 Características de Mycroft AI

Razón	Descripción
Código Fuente Abierto	El código de Mycroft AI puede ser inspeccionado, copiado, modificado y se puede contribuir de regreso, a diferencia de Alexa, Google Assistant, Cortana, Siri que son cajas negras, donde no se sabe su contenido, qué y cómo se hace.
Respeto de privacidad	La grabación de voz sólo funciona con permiso explícito, sin embargo, se puede deshabilitar en cualquier momento.
Múltiple hardware	Mycroft AI se ejecuta sobre una amplia gama de plataformas de hardware y software, RPI, Android, Linux PC.
Ligero	Diseñado para ejecutarse en hardware de bajo costo y bajo consumo de recursos y alimentación.
Comunidad	Comunidad vibrante, comprometida y servicial.

Elaboración propia

De la Tabla 8 se optó por evaluar las soluciones puntero, tanto Mycroft AI (Montgomery, Joshua, 2015) como Leon (Grenard, Louis, 2019) por cumplir precisamente con ser soluciones disponibles a la mano, instalación sobre Linux PC, documentación accesible y ejecución sencilla.

Sin embargo, ante problemas en el proceso de instalación, la solución Leon no provee documentación suficiente para la resolución de problemas con las dependencias con herramientas por ejemplo bibliotecas de *Node.js*, en su lugar la solución Mycroft AI sí cuenta dentro de su documentación con una sección

dedicada para problemas en la instalación o troubleshooting, es por esta razón que Mycroft AI se convirtió en la solución a emplear para el trabajo de tesis.

3.2.2. Instalación de Mycroft AI en Linux PC

Se decidió emplear una PC la cual a diferencia de una Raspberry PI cuenta con interfaces de audio, red y pantalla integradas, 2GB de RAM, Procesador Intel Celeron de 2 Núcleos con Sistema Operativo Lubuntu cumpliendo con los prerequisites para la correcta instalación y ejecución de Mycroft AI.

Dentro de las diferentes maneras de instalar Mycroft AI, se optó por instalarlo mediante el comando *git clone*, esto es, clonar el repositorio en github que lleva por nombre mycroft-core dentro del sistema local en el directorio home, y posteriormente, instalar las dependencias dentro del folder mycroft-core, los pasos a seguir son los siguientes:

```
cd ~/
git clone https://github.com/MycroftAI/mycroft-core.git
cd mycroft-core
bash dev_setup.sh
```

Además de instalar dependencias de herramientas, el script *dev_setup.sh* también instala y configura *virtualenv*, una herramienta para crear un ambiente aislado de Python 3, el cual se usa para trabajar de manera segura e independiente.

3.2.3. Ejecución de Mycroft

Dos scripts son los que permiten iniciar y parar de manera efectiva todos los servicios de Mycroft AI a través del ambiente *virtualenv*, estos son *start-mycroft.sh* y *stop-mycroft.sh*. Mientras que *start-mycroft.sh* tiene diversas posibilidades tales como iniciar todos los servicios de audio, voz, bus y skills o iniciar cada servicio por separado, comenzar interfaz de línea de comandos o comenzar sesión de debug, etc., el script *stop-mycroft.sh* detiene todo lo que se haya comenzado indiscriminadamente.

El comando y la salida en consola que inicia todos los servicios de Mycroft AI es:

```
$ ./start-mycroft.sh all
Starting all mycroft-core services
Initializing...
Starting background service bus
Starting background service skills
Starting background service audio
Starting background service voice
```

Inmediatamente después, para visualizar la interfaz de línea de comandos propia de Mycroft AI se emplea:

```
$ ./start-mycroft.sh cli
```

El resultado se muestra en la Figura 12.

Accediendo con los datos de la cuenta de Mycroft AI al enlace <https://sso.mycroft.ai/home.mycroft.ai/pair>, la Figura 13 muestra la configuración donde el código de registro de 6 dígitos lanzado por la consola de Linux en el dispositivo Mycroft AI se debe ingresar, adicionalmente parámetros generales como ubicación geográfica, tipo de voz y wake word default.

1 Device 2 Done!

Configure your new device

Pairing Code *

Code spoken by device

Name * Placement

Must be unique e.g. Kitchen, Bedroom, Office

Geographical Location

Country Region

Australia Northern Territory

City Time Zone

Darwin Australia/Darwin

Voice

British Male American Female American Male Google Voice

Wake Word

Hey Mycroft Christopher Hey Ezra Hey Jarvis

NEXT

Figura 13 Ventana de emparejamiento para dispositivo de Mycroft AI

Elaboración propia

Al terminar de ingresar la información correspondiente, la cuenta de Mycroft AI para el dispositivo Linux PC se muestra la información de configuración en la Figura 14, Figura 15, Figura 16 y Figura 17.

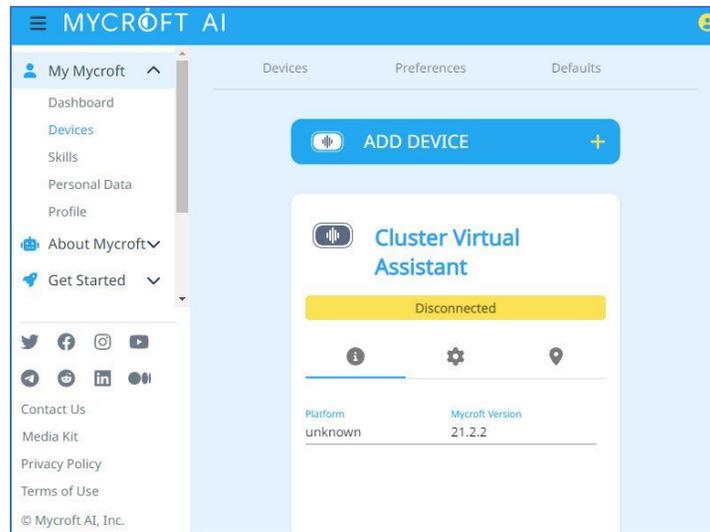


Figura 14 Nombre, Estado, Plataforma y Versión del dispositivo Mycroft AI

Elaboración propia

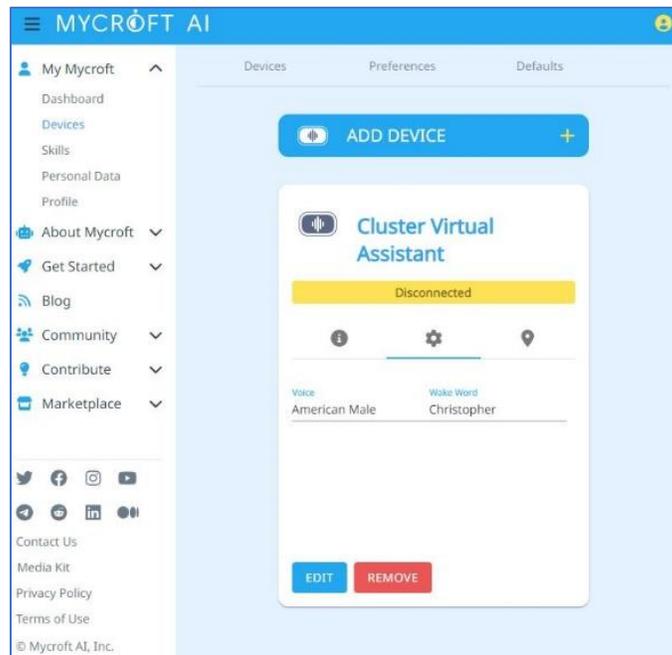


Figura 15 Tipo de Voz y Wake Word del dispositivo Mycroft AI

Elaboración propia

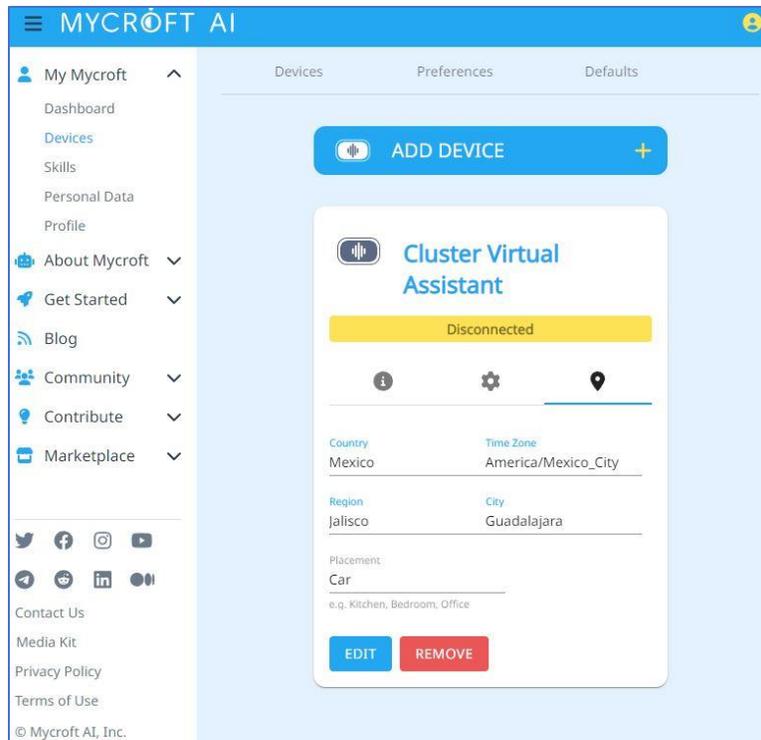


Figura 16 Ubicación geográfica y física del dispositivo Mycroft AI

Elaboración propia

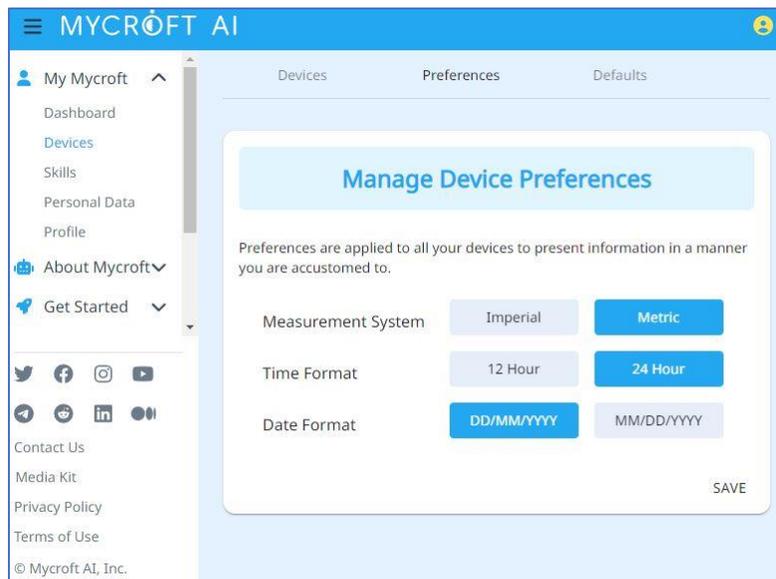


Figura 17 Preferencias, medidas, huso horario y fechas del dispositivo Mycroft AI

Elaboración propia

3.2.5. Mycroft modular

Mycroft AI es un conjunto de componentes modulares de software y hardware que emplean técnicas NLP y ML para proveer un asistente virtual de fuente abierta. Cada componente de la Tabla 7 puede ser personalizado, iniciado y detenido de manera independiente en Mycroft AI, a continuación, algunos detalles.

3.2.5.1 Detección de Wake Word

En el esquema tradicional de pregunta-respuesta, la *Wake Word* es una frase usada para indicar a Mycroft AI una orden o comando, por ejemplo “*Hey Mycroft*” es la palabra default, la cual puede ser personalizada a través de (University, Carnegie Mellon, 2014).

El software base empleado para la modificación de la *Wake Word* puede ser:

1. *PocketSphinx* (University, Pocket Sphinx, 2014). Es parte de un paquete más grande llamado CMUSphinx (University, CMU Sphinx, 2014) desarrollado por la Universidad de Carnegie Melon originalmente SPHINX-II (Huang, Alleva, Hwang, & Rosenfeld, 1993). Es un motor de reconocimiento de voz ligero con enfoque a dispositivos móviles y basado en la tecnología STT brinda acústica unificada y modelado de lenguaje a través de algoritmos como normalized feature representations, multiple-codebook semi-continuous hidden Markov models, between-word senones, y multi-pass search.
2. *Precise*. A diferencia de *PocketSphinx*, es una red neuronal entrenada en audio, aunque brinda mayor precisión, requiere ser entrenada.

Debido a que *PocketSphinx* emplea patrones de palabras y *Precise* emplea patrones de sonidos, por practicidad y simplicidad de personalización se optó por usar *PocketSphinx*.

Como primera adaptación del asistente virtual alineado al enfoque de UX del trabajo de tesis y a manera de alusión al nombre corto del patrocinador del programa de maestría Continental, se optó por dar de alta la nueva *Wake Word*

“Conti”, de hecho, sitios como (Heller, 2023) indican que la frecuencia de aparición de la palabra más parecida que es el verbo “continue”, es de 0.03%.

La adición de la nueva Wake Word se logró a través de dar de alta un nuevo contenedor de configuración dentro del apartado hotwords del archivo *mycroft.conf*, para la herramienta PocketSphinx se emplearon los fonemas del Carnegie Mellon University Dictionary (University, Carnegie Mellon, 2014). La palabra “CONTI” produce los correspondientes fonemas “K AA N T IY”, la Tabla 10 muestra la configuración listener la wake_word de nueva creación.

Tabla 10 Configuración de listener para la wake word Conti del dispositivo Mycroft AI

Contenedor de configuración	Descripción
<pre>"hotwords": { "conti": { "module": "pocketsphinx", "phonemes": "K AA N T IY", "threshold": 1e-40 } }</pre>	Nueva hotword de nombre conti que emplea pocketsphinx con los fonemas K AA N T IY con un umbral de sensibilidad de 1e-40.
<pre>"listener": { "wake_word": "conti" }</pre>	Nuevo listener agrega la wake word conti.

Elaboración propia

La Wake Word default “Hey Mycroft” se conservó, así ambas Wake Words son válidas para introducir comandos de voz.

3.2.5.2 Speech To Text

Mycroft AI emplea *Google STT* (Google, 2013) como motor default para la traducción de palabras habladas a frases de texto, el progreso en Deep Learning en transcripción de voz hace que modelos como LSTM RNN (Beaufays, Françoise, 2015) tengan un desempeño notable en reconocimiento de voz y transcripción de texto.

De manera independiente Mycroft trabaja con Mozilla para construir *DeepSpeech*, un motor completamente de fuente abierta. También soporta otras soluciones tales como *IBM Watson STT* y *wit.ai STT*.

3.2.5.3 Intent Parser

La salida del componente STT es procesada para identificar intenciones de ordenes o comandos, *Adapt* (Mycroft AI, 2020) es el intent parser default usado por Mycroft AI, sin embargo, es posible cambiar a *Padatious* que está basado en una red neuronal para mayor precisión a costa de entrenamiento.

3.2.5.4 Text to Speech

Mycroft AI convierte texto a audio con *Mimic* (Mycroft AI, 2022) como la herramienta default creada por Mycroft en alianza con VocaliID, Mimic está basada en el motor de síntesis de voz con texto de fuente abierta llamada CMU's Flite, algoritmos como Classification And Regression Tree y Finite State Transducer son empleados. Otras opciones son Mimic2 o Google TTS (Google, 2013).

3.2.5.5 Mycroft Skills

Mycroft AI cuenta con aplicaciones llamadas skills para ejecutar diversas funciones como temporizadores, alarmas, consulta de clima, tiempo, fecha, hora, etc., más aún ofrece la posibilidad de desarrollar skills adicionales a través de la herramienta basada en Python 3 *Mycroft Skills Kit (MSK)*.

3.3. CREAR LA FUNCIONALIDAD DE PREALIMENTACIÓN PARA ASISTENTE VIRTUAL

Como se mencionó previamente, las Mycroft Skills son aplicaciones de voz que proveen diversas funcionalidades a los usuarios, información, noticias, reproducción de medios, dispositivos y casas inteligentes, juegos, etc.

Los requisitos técnicos para crear y trabajar una Mycroft Skill son:

- a. Conocimiento de Python ya que las Skills están creadas con él.

- b. Cuenta de Github para trabajar con repositorios de Mycroft AI.
- c. Un hardware donde se instale y ejecute Mycroft AI.

La creación de una Skill, implica definir la acción que realizará y cómo lo utilizarán los usuarios. Una buena Skill cumple con una o más necesidades de los usuarios, es por esto que una Skill es el medio idóneo que da sentido y ayuda a materializar el objetivo general de este trabajo de tesis.

Mycroft AI opera con un enfoque basado en Diseño Centrado en el Usuario (User Centered Design) ó Design Thinking (Dam, 2021), el cual es una metodología que provee una solución que resuelve necesidades frecuentemente complejas de un humano, para lograrlo, el proceso de Design Thinking consiste en las etapas que a continuación se describen:

1. *Empatizar* - Investigar las necesidades de usuarios
2. *Definir/Diseñar* - Declarar las necesidades y problemas de los usuarios
3. *Idear* - Ideas y desafiar suposiciones.
4. *Prototipar* - Comenzar a crear soluciones.
5. *Probar* – Poner a prueba las soluciones.

Finalmente repetir reiteradamente hasta que se produzca el resultado esperado.

Para este trabajo de tesis se identificaron tres tipos de necesidades de usuario respecto a los telltales del IP manejado por asistente virtual, la Tabla 11 resume esta información.

Tabla 11 Necesidades, casos de uso y ejemplos para los telltales de un IP manejado por asistente virtual

Necesidad	Caso de uso	Ejemplo
Ser provisto de notificaciones de estado o cambio de estado de un telltale.	El usuario es notificado y decide las acciones a tomar con la información con la que él mismo cuenta.	Ante bajo nivel de combustible es responsabilidad del usuario conocer la ubicación de la estación de servicio y dirigirse a ella.
Ser provisto de sugerencias ante un cambio de estado de telltale.	El usuario decide las acciones a tomar con la información con la que cuenta el asistente virtual.	Ante bajo nivel de combustible, brindar instrucciones para llegar a la estación de servicio más cercana, sigue siendo responsabilidad del usuario dirigirse a ella.
Ser provisto de acciones ante un cambio de estado de telltale.	El asistente virtual con base a la información con la que cuenta decide la acción a tomar.	Ante bajo nivel de combustible, el asistente virtual con información de su entorno conoce e instruye a un auto autónomo a dirigirse a la estación de servicio más cercana.

Elaboración propia

El enfoque de Design Thinking sintetiza el manejo de un problema de la siguiente manera: **Cuando** _____, **Yo quiero** _____, **de tal manera que puedo** _____.

Las tres necesidades del proyecto de tesis se alinean a la perfección con este enfoque, de esta manera, para el indicador vital Cinturón de Seguridad, tenemos:

Cuando

el Cinturón de seguridad no esté abrochado,

Yo quiero

escuchar una notificación ó

recibir una sugerencia amable ó

ser sujetado automáticamente al asiento,

de tal manera que puedo

lograr obtener seguridad en mi viaje.

Todos y cada uno de los indicadores virtuales con los que el asistente virtual trabaja la prealimentación emplean este enfoque.

A manera de resumen, dentro de Mycroft AI, el *Flujo de una Skill*, contempla:

- a. Qué palabras hablará el usuario.
- b. Qué respuesta hablará Mycroft AI.
- c. Qué datos requerirá la Skill para funcionar.
- d. Qué dependencias o paquetes adicionales se requerirán.

3.3.1. Mycroft Skill Kit

La MSK es una utilidad instalada con Mycroft AI que facilita la creación, carga (upload) y actualización (upgrade) de Skills. El comando para ejecutar MSK en terminal desde el directorio mycroft-core es *mycroft-msk*.

Si se ejecuta sin parámetros, arroja documentación sobre su uso. *mycroft-msk create*, ejecuta un script interactivo que lanza unas pocas preguntas con la finalidad de generar una plantilla de Skill, su objetivo principal es atender el *Flujo de una Skill*.

La Tabla 12 muestra las preguntas del script y las respuestas que le fueron dadas al ejecutar el comando *mycroft-msk create*.

Tabla 12 Script de MSK para la skill Telltales

Pregunta de script	Respuesta
Nombre de Skill (Por lo regular debajo de 22 caracteres)	Telltales
Comandos (Utterances) del usuario a Mycroft AI	What is the car seatbelt status
Diálogo de respuesta de la Skill de Mycroft AI al usuario	Car seatbelt is fastened
Descripción corta (1 línea, máximo 40 caracteres)	Telltale status / Indicator level retrieval
Descripción larga	Skill to retrieve instrument panel's current telltale status or indicator level
Autor (Nombre / Github user)	ChaklKan
Categoría a la que pertenece en el Mycroft Skills Marketplace	IoT
Tags para identificación rápida de la Skill	Ninguna

Elaboración propia

Al terminar de ingresar la información, Mycroft AI pregunta si se desea crear un repositorio de GitHub para la Skill, sólo será requerido si se desea publicar en el Marketplace. Para el trabajo de tesis se optó por no crear ningún repositorio y mantener el trabajo de manera local.

El resultado de los pasos anteriores es la creación de la **Skill Telltales** en el directorio /opt/mycroft/skills Figura 18.

```
chakikkan@lubuntuchak: /opt/mycroft/skills/telltales-skill$ ls -l
total 24
-rw-rw-r-- 1 chakikkan chakikkan 2091 Dec  6 21:28 __init__.py
drwxrwxr-x 3 chakikkan chakikkan 4096 Nov  8 21:17 locale
-rw-rw-r-- 1 chakikkan chakikkan 1009 Nov  8 21:21 manifest.yml
drwxrwxr-x 2 chakikkan chakikkan 4096 Dec  6 21:28 __pycache__
-rw-rw-r-- 1 chakikkan chakikkan  450 Nov  8 21:26 README.md
-rw-rw-r-- 1 chakikkan chakikkan  631 Nov  8 21:21 settingsmeta.yaml
```

Figura 18 Directorio de skill Telltales

Elaboración propia

3.3.2. Directorio Local

El directorio *locale* Figura 19 contiene subdirectorios con archivos para cada lenguaje que se soporta. Aunque Mycroft AI soporta diversos lenguajes, inglés es el predefinido por su universalidad, por lo que se encuentra un directorio *en-us* para identificarlo.

```
chakikkan@LubuntuChak:/opt/mycroft/skills/telltales-skill$ cd locale/en-us/
chakikkan@LubuntuChak:/opt/mycroft/skills/telltales-skill/locale/en-us$ ls -l
total 12
-rw-rw-r-- 1 chakikkan chakikkan 60 Dec 6 21:05 telltales.dialog
-rw-rw-r-- 1 chakikkan chakikkan 35 Dec 6 21:26 telltales.entity
-rw-rw-r-- 1 chakikkan chakikkan 161 Dec 6 21:08 telltales.intent
```

Figura 19 Directorio locale de skill Telltales

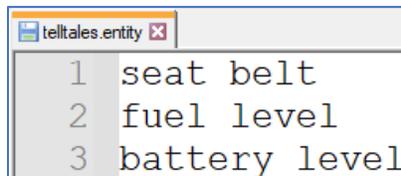
Elaboración propia

Los archivos **.intent* y **.dialog* contienen las frases creadas mediante el MSK, correspondientemente, los comandos hacia la Skill y las respuestas de la Skill, un dato curioso es que las respuestas contenidas en **.dialog* son elegidas de manera aleatoria a cualquier comando, esto para dar la impresión de una interacción más natural.

Los archivos **.entity* son creados manualmente para proveer de flexibilidad a los **.intent*, una cadena de caracteres comodín o wild-card encerrada entre llaves es empleada para manejar diferentes tipos de entradas en las frases, estas entradas son reenviadas a la Skill para su correspondiente procesamiento.

3.3.3. Archivos intent, entity y dialog

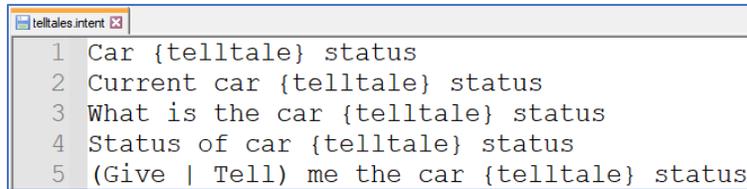
Los archivos **.intent* Figura 21, **.dialog* Figura 22 y **.entity* Figura 20 fueron modificados manualmente de acuerdo a las necesidades de la Skill Telltales.



```
telltales.entity
1 seat belt
2 fuel level
3 battery level
```

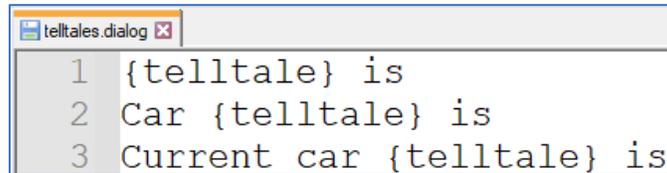
Figura 20 Archivo telltales.entity

Elaboración propia



```
1 Car {telItale} status
2 Current car {telItale} status
3 What is the car {telItale} status
4 Status of car {telItale} status
5 (Give | Tell) me the car {telItale} status
```

Figura 21 Archivo telItales.intent
Elaboración propia



```
1 {telItale} is
2 Car {telItale} is
3 Current car {telItale} is
```

Figura 22 Archivo telItales.dialog
Elaboración propia

Un caso de uso para estos tres archivos es el siguiente:

Comando: *Car seat belt status*

Diálogo de respuesta: *Current car seat belt is*

Los diálogos de respuesta se dejaron intencionalmente incompletos, ya que la Skill maneja mediante variables (atributos de clase), los correspondientes estados de cada uno de los telItales.

3.3.3.1 Modos de prealimentación

La funcionalidad de prealimentación se plantea trabajar en tres diferentes modos:

1. **Modo Notificación** – Se limita a reportar estado de los telItales.
2. **Modo Sugerencia** – Además de reporta el estado de los telItales, con base al contexto del auto, se brinda información adicional y sugerencias de acción a tomar por parte del conductor o pasajero.
3. **Modo Acción** - Además de reporta el estado de los telItales, con base al contexto del auto, se toman las acciones pertinentes sin interacción esperada por parte del conductor o pasajero, con la finalidad de resolver una situación de eventual falla.

La Figura 23 muestra el diagrama de estados correspondiente a los modos de prealimentación del sistema.

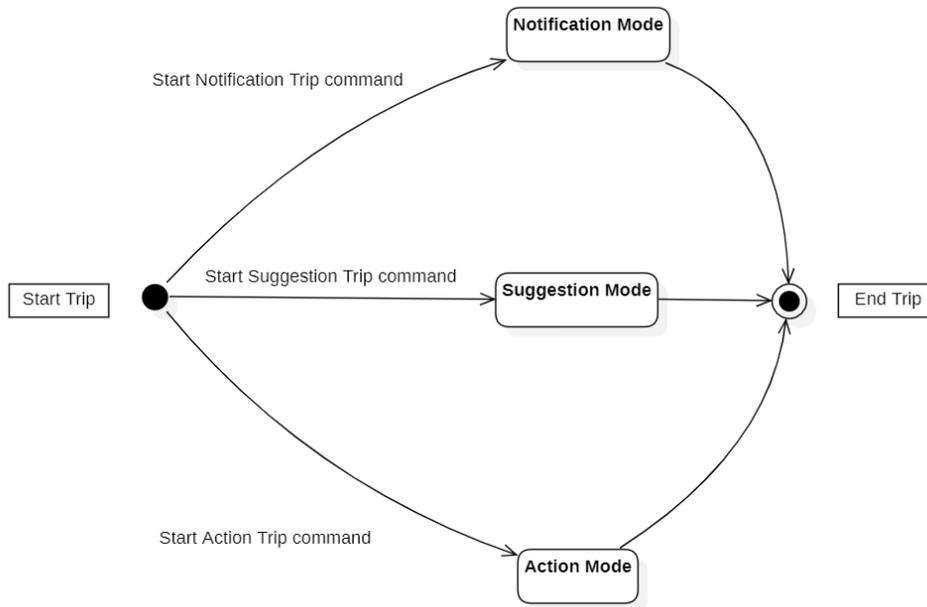


Figura 23 Diagrama de estados o modos de prealimentación

Elaboración propia

Naturalmente tres archivos *.intent son los requeridos para atender precisamente a cada una de las transiciones para ingresar a los correspondientes modos de prealimentación, estos archivos se muestran en la Figura 24, Figura 25 y Figura 26.

```
trip_notification.intent x
1 start notification trip
```

Figura 24 Archivo trip_notification.intent

Elaboración propia

```
trip_suggestion.intent x
1 start suggestion trip
```

Figura 25 Archivo trip_suggestion.intent

Elaboración propia

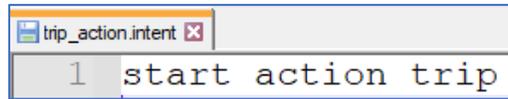


Figura 26 Archivo trip_action.intent

Elaboración propia

El comportamiento dinámico se muestra en la Figura 27, Figura 28 y Figura 29 correspondientemente.

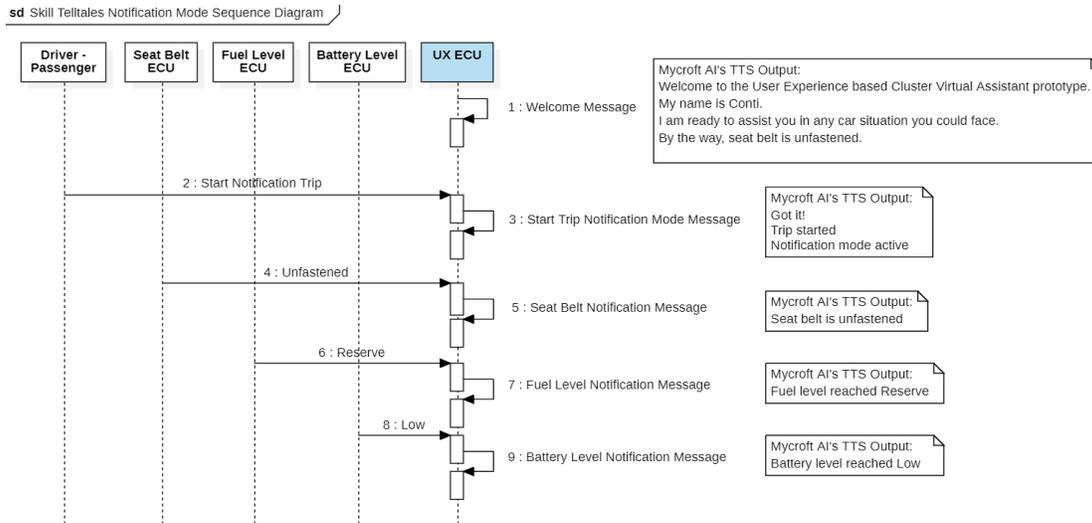


Figura 27 Secuencia de prealimentación en modo notificación

Elaboración propia

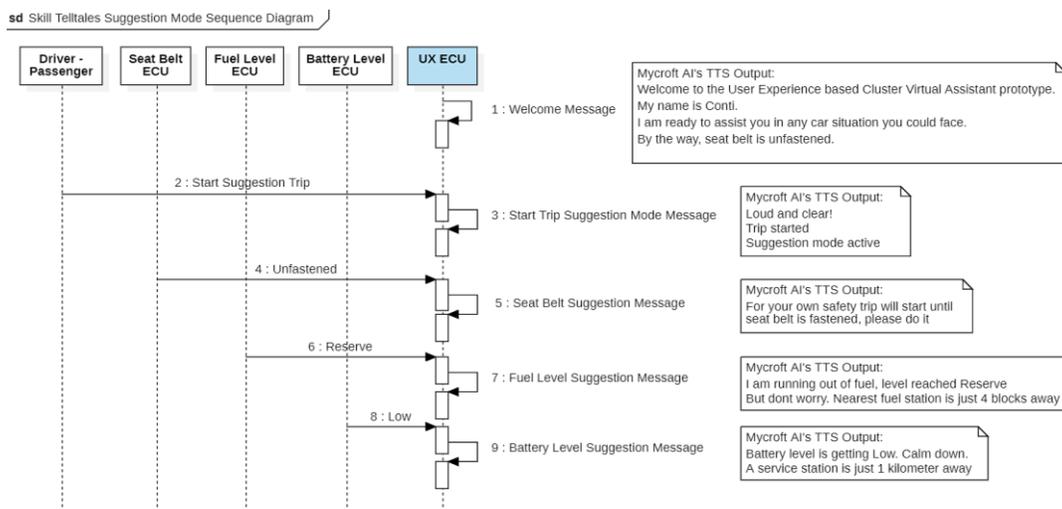


Figura 28 Secuencia de prealimentación en modo sugerencia

Elaboración propia

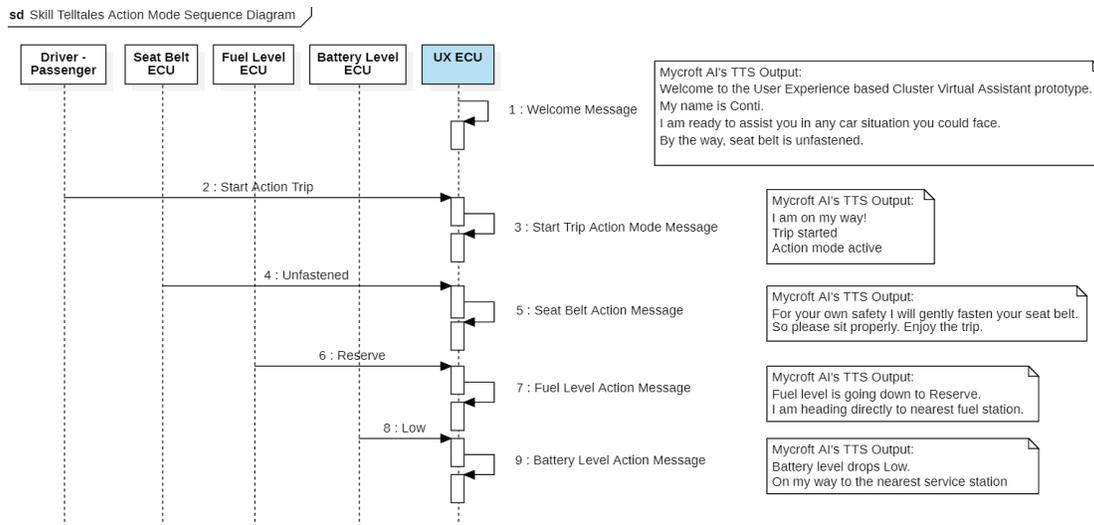


Figura 29 Secuencia de prealimentación en modo acción

Elaboración propia

3.3.4. Archivo `__init__.py`

Este archivo es el núcleo en Python de la Skill donde se realizan las acciones como:

3.3.4.1 Importar bibliotecas requeridas

```
from mycroft import MycroftSkill, intent_handler
import time
```

3.3.4.2 Definir clase que hereda de MycroftSkill

```
class Telltales(MycroftSkill):
```

3.3.4.3 Definir métodos para manejo de intents y diálogos

3.3.4.3.1 `__init__`

Constructor, declara variables o acciones de configuración. Sin acceso a métodos de la clase MycroftSkill debido a que no hay instancia de la clase.

```
def __init__(self):
    MycroftSkill.__init__(self)
```

3.3.4.3.2 Initialize

Configuración final de la Skill con acceso a métodos de la clase MycroftSkill, ya que la Skill ha sido construida y registrada en el sistema.

```
def initialize(self):
    self.register_entity_file('telltales.entity')
#entity file with matching wildcard for telltales must be registered

    self.log.info(" Telltales Skill has been successfully loaded, bravoooo! ")
#Log skill init message

    self.speak_dialog("Welcome to the User Experience based Cluster Virtual
Assistant prototype.")

    self.speak_dialog("My name is Conti.")

    self.speak_dialog("I am ready to assist you in any car situation you could
face.")

    self.speak_dialog("By the way, seat belt is unfastened.")
```

3.3.4.3.3 Handle_telltales (intent_handler)

Método que maneja solicitudes genéricas de información de telltales.

```
@intent_handler('telltales.intent') #decorator for intent handler
def handle_telltales(self, message): #intent handler for telltales
    telltale = message.data.get('telltale') #get wildcard word from message
    self.speak_dialog('telltales', data={ #speak wildcard
        'telltale': telltale
    })
#switch case for telltales dialogs
```

```
match telltale:
    case "seat belt":
        self.speak_dialog('Unfastened')
    case "fuel level":
        self.speak_dialog('Reserve')
    case "battery level":
        self.speak_dialog('Low')
```

3.3.4.3.4 Handle_notification (intent_handler)

Método que maneja la prealimentación en modo notificación.

```
@intent_handler('trip_notification.intent') #decorator for intent handler
def handle_notification(self, message): #intent handler for notification trip
    self.speak_dialog('Got it!')
    self.speak_dialog('Trip started')
    self.speak_dialog('Notification mode active')
    time.sleep(5)
    self.speak_dialog('Seat belt is unfastened')
    time.sleep(5)
    self.speak_dialog('Fuel level reached Reserve')
    time.sleep(5)
    self.speak_dialog('Battery level reached Low')
```

3.3.4.3.5 Handle_suggestion (intent_handler)

Método que maneja la prealimentación en modo sugerencia.

```
@intent_handler('trip_suggestion.intent') #decorator for intent handler
def handle_suggestion(self, message): #intent handler for suggestion trip
    self.speak_dialog('Loud and clear!')
    self.speak_dialog('Trip started')
    self.speak_dialog('Suggestion mode active')
```

```

time.sleep(8)
self.speak_dialog('For your own safety trip will start until seat belt is
fastened, please do it')
time.sleep(10)
self.speak_dialog('I am running out of fuel, level reached Reserve')
self.speak_dialog('But dont worry')
self.speak_dialog('Nearest fuel station is just 4 blocks away')
time.sleep(10)
self.speak_dialog('Battery level is getting Low')
self.speak_dialog('Calm down')
self.speak_dialog('A service station is just 1 kilometer away')

```

3.3.4.3.6 Handle_action (intent_handler)

Método que maneja la prealimentación en modo acción.

```

@intent_handler('trip_action.intent') #decorator for intent handler
def handle_action(self, message): #intent handler for action trip
    self.speak_dialog('I am on my way!')
    self.speak_dialog('Trip started')
    self.speak_dialog('Action mode active')
    time.sleep(8)
    self.speak_dialog('For your own safety I will gently fasten your seat belt')
    self.speak_dialog('So please sit properly')
    self.speak_dialog('Enjoy the trip')
    time.sleep(10)
    self.speak_dialog('Fuel level is going down to Reserve')
    self.speak_dialog('I am heading directly to nearest fuel station')
    time.sleep(10)
    self.speak_dialog('Battery level drops Low')
    self.speak_dialog('On my way to the nearest service station')

```

3.3.4.4 Creación de la instancia de la Skill

create_skill – el código final que crea la instancia de la Skill y que Mycroft AI es responsable de cargar.

```
def create_skill():  
    return Telltales()
```

3.3.5. Archivo README.md

Contiene información entendible por humanos acerca de la Skill, esta información proviene del MSK, tal como descripción corta, descripción larga, autor, categoría de la skill.

```
#  src="https://raw.githubusercontent.com/FortAwesome/Font-Awesome/master/svg/solid/robot.svg" card_color="#22A7F0" width="50" height="50" style="vertical-align: bottom;"/> Telltales  
Telltale status / Indicator level retrieval
```

```
## About
```

```
Skill to retrieve instrument panel's current telltale status or indicator level
```

```
## Examples
```

```
* "What is the {telltale} status"
```

```
* "Current {telltale} status"
```

```
## Credits
```

```
ChaklKan
```

```
## Category
```

```
**Transport**
```

```
IoT
```

```
## Tags
```

3.3.6. Archivo settingsmeta.yaml

Contiene la configuración parametrizada de la Skill almacenada en el perfil de Mycroft AI <https://sso.mycroft.ai/home.mycroft.ai>, fecha, hora, unidades de tiempo, unidades de medida, ubicación, tipo de voz, etc. Para la Telltales Skill no se requiere modificar ningún parámetro, de tal manera que no se altera este archivo.

3.3.7. Archivo manifest.yml

Contiene todas las dependencias externas de paquetes requeridos por la Skill. Para la Telltales Skill no se definieron dependencias externas.

3.4. CASOS DE PRUEBA

Los casos de prueba para el módulo de voz se encuentran en la Tabla 13 .

Tabla 13 Casos de Prueba de Skill Telltales

Intent, Utterance, Comando	Diálogo
Car seat belt status	Current car seat belt is Unfastened
What is the car fuel level status	Fuel level status is Reserve
Give me the car battery level status	Car battery level status is Low
Start Notification Trip	Got it! Trip started Notification mode active (5 seconds artificial delay) Seat belt is unfastened (5 seconds artificial delay) Fuel level reached Reserve (5 seconds artificial delay) Battery level reached Low
Start Suggestion Trip	Loud and clear! Trip started Suggestion mode active (8 seconds artificial delay) For your own safety trip will start until seat belt is fastened, please do it (10 seconds artificial delay) I am running out of fuel, level reached Reserve But dont worry Nearest fuel station is just 4 blocks away (10 seconds artificial delay) Battery level is getting Low Calm down A service station is just 1 kilometer away
Start Action Trip	I am on my way! Trip started Action mode active (8 seconds artificial delay) For your own safety I will gently fasten your seat belt So please sit properly Enjoy the trip (10 seconds artificial delay) Fuel level is going down to Reserve I am heading directly to nearest fuel station (10 seconds artificial delay) Battery level drops Low On my way to the nearest service station

Elaboración propia

4. RESULTADOS

Los resultados logrados se enlistan a continuación:

- a. Resumen de investigación sobre soluciones actuales en la industria automotriz de productos de asistencia virtual basada en voz, las cuales confirman que es una tendencia en incremento, la omnipresencia dentro de los automóviles actuales y futuros a corto plazo también es prueba suficiente de su relevancia como un elemento clave de gran valor al que se le debe poner especial atención.
- b. Resumen de investigación sobre soluciones contemporáneas de asistentes virtuales de código abierto, que evidencian las fuertes motivaciones a nivel mundial para trabajar, contribuir a la mejora y refinamiento de los métodos, algoritmos y técnicas relacionadas al NLP.
- c. Mycroft AI mostró ser la solución de asistente virtual de código abierto más completa, su instalación y configuración en una PC con Linux es sencilla, tanto la documentación y sección dedicada a problemas está disponible y es útil, finalmente, la creación y puesta en marcha de aplicaciones de voz o skills, es directa con ayuda del MSK.
- d. La skill Telltales creada siguiendo el enfoque UCD y el método de Design Thinking sobre los tres indicadores vitales cinturón de seguridad, nivel de combustible y nivel de batería, representa concretamente el módulo de comunicación por voz de un panel de instrumentos automotriz, el cual posee las siguientes características:
 - a. Simple, debido a que el folder de salida y la dependencia de archivos son reducidas y sencillas de manejar.
 - b. Portable, debido a que puede ser empleado en hardware como RPI, Android o Linux que tenga a Mycroft AI instalado.
 - c. Personalizable, a través de la adaptación de los archivos intent, dialog, entity, así como la clase de la skill basada en Python.
- e. Finalmente, se mostró en acción tanto la característica de retroalimentación empleando el esquema tradicional de pregunta-respuesta sobre el estado de los indicadores vitales, la grabación puede consultarse en (Hernandez,

Mycroft Skill Telltales v1.0, 2023), como la característica de prealimentación la cual mediante en un ambiente simulado de un viaje en tres diferentes modos, Notificación, Sugerencia y Acción, diferentes ECUs simulados envían mensajes de manera automática al módulo de comunicación por voz y dependiendo del modo seleccionado es el correspondiente mensaje de salida prealimentado al conductor y/o pasajero, la grabación puede consultarse en (Hernandez, Mycroft Skill Telltales v2.0, 2023).

CONCLUSIONES

El resultado del análisis e investigación del estado del arte de la asistencia virtual en los distintos rubros tales como academia, investigación y comercial, así como el resultado de investigación de soluciones contemporáneas de asistentes virtuales de código abierto, permitieron justificar la importancia de este trabajo así como identificar la solución más completa de asistente virtual de código abierto para materializar un prototipo que bien se puede considerar bien como un sistema base o bien como un módulo de comunicación de voz integrable en un sistema de mayor complejidad que opera indicadores vitales propios de un panel de instrumentos automotriz empleando principalmente la prealimentación como técnica de diseño de experiencia de usuario extendiendo la tradicional técnica de retroalimentación.

RECOMENDACIONES

El alcance de este trabajo se limitó a un conjunto reducido de indicadores vitales de un panel de instrumentos automotriz, una lista corta de comandos, entidades y diálogos se introdujo en los archivos correspondientes para mantener la complejidad simple y manejable, sin embargo, en la realidad la cantidad de indicadores son numerosos tanto los relevantes para la seguridad, como los no relevantes para la seguridad, es por esta razón que una recomendación es incrementar el soporte a una cantidad mayor de indicadores sin perder de vista el enfoque UCD así como la metodología Design Thinking.

A manera de sugerencia para trabajo futuro, se identifican vastas posibilidades de conexión y expansión del módulo de comunicación de voz presentado en este trabajo, a través del uso de innumerables bibliotecas de Python, por ejemplo para importar cualquier tipo de información, en cualquier formato o fuente, información proveniente de diferentes sistemas de interconexión a través de buses de comunicación y tecnologías disponibles en el dominio automotriz tales como CAN, LIN, ETH, MOST, GPS, V2X, entre otras, esto con la finalidad de procesar y presentar información relevante al usuario final.

APORTACIÓN DE LA TESIS

La principal contribución de este trabajo es sentar los fundamentos de la evolución de los paneles de instrumentos automotrices contemporáneos, desde sus orígenes como un mero dispositivo físico hacia un dispositivo virtual completo, esto es, los usuarios de automóviles demandan de vehículos del futuro la asistencia virtual basada en voz, lo cual para los indicadores vitales se logró mediante el uso de Mycroft AI, además la característica única con la que cuenta, el diseño y desarrollo de aplicaciones basadas en el enfoque UCD y el método Design Thinking marcan la diferencia ante competidores.

APORTACIÓN SOCIAL DE LA TESIS

Dentro de la industria automotriz, las empresas grandes, medianas, pequeñas y hasta start-ups son las principales interesadas y beneficiarias de este trabajo, debido a que los módulos de comunicación de voz son completamente compatibles con cualquier tipo de productos HMI dentro del auto. El hecho que Mycroft AI sea una solución creada con código abierto, por una parte, hace que su uso otorgue ventajas tecnológicas y económicas a sus beneficiarios, más aún y desde mi punto de vista el más importante, es que promueve el espíritu de mantener en el dominio público las aportaciones en este campo del conocimiento humano, beneficiando directamente a la comunidad de código abierto internacional.

REFERENCIAS

- Aggarwal, V., & Binay, D. (2 de Abril de 2021). *Estados Unidos Patente n° US 2020/0175386 A1*.
- Arora, S., Athavale, V., Maggu, H., & Agarwal, A. (2020). Artificial Intelligence and Virtual Assistant—Working Model. *Mobile Radio Communications and 5G Networks* (págs. 163-171). Singapore: Springer.
- Beaufays, Françoise. (August de 2015). *The neural networks behind Google Voice transcription*. (Google) Recuperado el 31 de August de 2022, de <https://ai.googleblog.com/2015/08/the-neural-networks-behind-google-voice.html>
- Boeriu, Horatiu. (24 de September de 2018). *Microsoft CEO mentions the successful BMW integration of Alexa and Cortana*. (BMW Group) Recuperado el 17 de May de 2023, de BMW Blog: <https://www.bmwblog.com/2018/09/24/microsoft-ceo-mentions-the-successful-bmw-integration-of-alex-and-cortana/>
- Bright, Stan. (2022). *Mycroft-core Alternatives*. (LibHunt) Recuperado el 14 de Octubre de 2022, de Libhunt: <https://www.libhunt.com/r/mycroft-core>
- Cloud Google. (s.f.). *IA de Text-to-Speech*. (Google Inc) Recuperado el 20 de May de 2023, de Cloud Google: <https://cloud.google.com/text-to-speech?hl=es>
- Dam, R. F. (Junio de 2021). *The 5 Stages in the Design Thinking Process*. (Interaction Design Foundation) Recuperado el 14 de Octubre de 2022, de Interaction Design: <https://www.interaction-design.org/literature/article/5-stages-in-the-design-thinking-process>
- Ford. (2023). *Amazon Alexa & Ford® | Alexa Built-In with Ford® Streaming*. (Ford Motor Company) Recuperado el 17 de May de 2023, de Ford: <https://www.ford.com/alex/>
- Friedman, S., Remegio, P., Falkenmayer, T., Kyle, R., Kakimi, R., Heide, L., & Puranik, N. (10 de Octubre de 2019). *Estados Unidos Patente n° US 2019/0311241 A1*.
- Gazi Yalçın, O. (2021). *Applied Neural Networks with TensorFlow 2*. Berkeley, CA: Apress.
- General Motors. (October de 2016). *Hello, OnStar – Meet Watson*. (General Motors Company) Recuperado el 17 de May de 2023, de GM: <https://news.gm.com/newsroom.detail.html/Pages/news/us/en/2016/oct/1025-watson.html>
- Google. (13 de Noviembre de 2013). *Speech-to-Text*. (Google) Recuperado el 14 de Octubre de 2022, de Google Cloud: <https://cloud.google.com/speech-to-text>

- Government, U. S. (17 de August de 2005). *Federal Motor Vehicle Safety Standards; Controls, Telltales and Indicators*. (NHTSA) Recuperado el 04 de September de 2022, de Federal Register: <https://www.federalregister.gov/documents/2005/08/17/05-16325/federal-motor-vehicle-safety-standards-controls-telltales-and-indicators>
- Grenard, Louis. (2019). *Meet your virtual brain*. (Leon AI) Recuperado el 14 de Octubre de 2022, de Get Leon: <https://getleon.ai/>
- Heller, T. (2023). *Word Frequency in Present-Day English*. Obtenido de Thomas' Language Playground: <https://tmh.conlang.org/word-frequency/>
- Hernandez, R. (19 de January de 2023). *Mycroft Skill Telltales v1.0*. Obtenido de Youtube: <https://youtu.be/vKNKNcGa-eU>
- Hernandez, R. (24 de May de 2023). *Mycroft Skill Telltales v2.0*. Obtenido de Youtube: <https://www.youtube.com/watch?v=aysqF4OYaNE>
- Honda. (6 de January de 2017). *Honda Introduces "Cooperative Mobility Ecosystem" at CES 2017*. (Honda Global) Recuperado el 17 de May de 2023, de Honda: <https://global.honda/newsroom/news/2017/c170106eng.html>
- Huang, X., Allewa, F., Hwang, M.-Y., & Rosenfeld, R. (1993). An Overview of the SPHINX-II Speech Recognition System. *Proceedings of the workshop on Human Language Technology (HLT '93)*, 81-86.
- Hyundai. (s.f.). *Send destinations, start or select the temperature of your Hyundai using Google Home and Bluelink*. (Hyundai Motor America) Recuperado el 17 de May de 2023, de Hyundai: <https://web-owners.hyundaiusa.com/us/en/resources/blue-link/using-blue-link-with-google-home>
- Kraus, M., Wagner, N., Callejas, Z., & Minker, W. (2021). The Role of Trust in Proactive Conversational Assistants. *IEEE Access*, 9, 112821-112836.
- Lens. (20 de July de 2023). *Patent Analysis*. Obtenido de Lens.org: <https://www.lens.org/lens/search/patent/analysis>
- Majji, K., & Baskaran, K. (2021). Artificial Intelligence Analytics—Virtual Assistant in UAE Automotive Industry. *Inventive Systems and Control* (págs. 309-322). Singapore: Springer.
- Media, M.-B. G. (16 de December de 2016). *Mercedes-Benz delivers integration of the Google Assistant*. (Mercedes-Benz) Recuperado el 17 de May de 2023, de Mercedes-Benz Media: <https://media.mercedes-benz.com/article/42118a43-5c19-487b-883b-a11beac9d7cb>
- Mercedes-Benz Group Media. (16 de Enero de 2021). *Mercedes-Benz presenta MBUX Hyperscreen en CES con nuevas funciones inteligentes como*

- «Mercedes Travel Knowledge». (Mercedes-Benz) Recuperado el 16 de Agosto de 2022, de Mercedes-Benz Group Media: <https://cafrealvolante.com/2021/01/16/mercedes-benz-presenta-mbox-hyperscreen-en-ces-con-nuevas-funciones-inteligentes-como-mercedes-travel-knowledge/>
- Montgomery, Joshua. (2015). *Mycroft AI*. (Mycroft AI, Inc) Recuperado el 14 de Octubre de 2022, de Mycroft AI: <https://mycroft.ai/>
- Mousa, H. (19 de Diciembre de 2018). *10 Top Open-source Voice Assistants Projects for Developers*. (MEDevel) Recuperado el 14 de Octubre de 2022, de Medevel: <https://medevel.com/10-open-source-voice-assistants/>
- Mycroft AI. (2020). *Adapt*. (Mycroft) Recuperado el 20 de May de 2023, de Mycroft AI: <https://mycroft-ai.gitbook.io/docs/mycroft-technologies/adapt>
- Mycroft AI. (2020). *Why use Mycroft AI?* (Mycroft) Recuperado el 22 de Enero de 2023, de Mycroft AI: <https://mycroft-ai.gitbook.io/docs/about-mycroft-ai/why-use-mycroft>
- Mycroft AI. (2021). *Padatious*. (Mycroft) Recuperado el 20 de May de 2023, de Mycroft AI: <https://mycroft-ai.gitbook.io/docs/mycroft-technologies/padatious>
- Mycroft AI. (2022). *Mimic TTS*. (Mycroft) Recuperado el 20 de May de 2023, de Mycroft AI: <https://mycroft-ai.gitbook.io/docs/mycroft-technologies/mimic-tts>
- Nissan. (17 de April de 2017). *Nissan Intelligent Mobility: Nissan presenta su asistente virtual Cortana diseñado para facilitar la vida cotidiana de los conductores*. (Nissan) Recuperado el 17 de May de 2023, de Nissan: <https://mexico.nissannews.com/es-MX/releases/nissan-intelligent-mobility-nissan-presenta-su-asistente-virtual-cortana-dise-ado-para-facilitar-la-vida-cotidiana-de-los-conductores>
- Norman, D. (2013). *The Design of Everyday Things*. New York: Basic Books.
- Reid, K. (17 de Marzo de 2020). *Private By Design: Free and Private Voice Assistants*. (makezine) Recuperado el 14 de Octubre de 2022, de Makezine: <https://makezine.com/article/home/connected-home/private-by-design-free-and-private-voice-assistants/>
- Ribbens, W. B. (2017). Automotive Instrumentation. En W. B. Ribbens, *Understanding Automotive Electronics - An Engineering Perspective (8th Edition)* (pág. 412). Michigan, USA: Elsevier.
- Team, S. (October de 2017). *Hey Siri: An On-device DNN-powered Voice Trigger for Apple's Personal Assistant*. (Machine Learning Research) Recuperado el 31

de August de 2022, de <https://machinelearning.apple.com/research/hey-siri>

Toyota. (4 de January de 2017). *Toyota Concept-i Makes the Future of Mobility Human*. (Toyota Newsroom) Recuperado el 17 de May de 2023, de Toyota: <https://pressroom.toyota.com/toyota-concept-i-future-of-mobility-human-cs-2017/>

University, C. M. (25 de Marzo de 2014). *CMU Sphinx*. (Carnegie Mellon University) Recuperado el 14 de Octubre de 2022, de OPEN SOURCE SPEECH RECOGNITION TOOLKIT: <https://cmusphinx.github.io/>

University, C. M. (25 de Marzo de 2014). *Pocket Sphinx*. (Carnegie Mellon University) Recuperado el 14 de Octubre de 2022, de Github: <https://github.com/cmusphinx/pocketsphinx>

University, Carnegie Mellon. (20 de Noviembre de 2014). *The CMU Pronouncing Dictionary*. (Carnegie Mellon University) Recuperado el 14 de Octubre de 2022, de Speech CS CMU: <http://www.speech.cs.cmu.edu/cgi-bin/cmudict>

Yong, Z., Jinyu, L., Shixiong, Z., Liping, C., & Yifan, G. (2018). Domain and Speaker Adaptation for Cortana Speech Recognition. *Speech and Signal Processing (ICASSP)* (págs. 5984-5988). Calgary, AB, Canada: IEEE International Conference on Acoustics. doi:10.1109/ICASSP.2018.8461553

Young-Bum, K. (07 de June de 2018). *The Scalable Neural Architecture behind Alexa's Ability to Select Skills*. (Amazon Science) Recuperado el 31 de August de 2022, de Amazon Science: <https://www.amazon.science/blog/the-scalable-neural-architecture-behind-alexas-ability-to-select-skills>

YourTechDiet. (2022). *List of Top 7 Open Source Voice Assistants*. (YourTechDiet) Recuperado el 14 de Octubre de 2022, de YourTechDiet: <https://yourtechdiet.com/blogs/open-source-voice-assistants/>