



UNIVERSIDAD TÉCNICA DE BABAHOYO

FACULTAD DE ADMINISTRACIÓN, FINANZAS E INFORMÁTICA

PROCESO DE TITULACIÓN

JUNIO 2021 – NOVIEMBRE 2021

**EXAMEN COMPLEXIVO DE GRADO O DE FIN DE CARRERA PRUEBA
PRÁCTICA**

INGENIERÍA EN SISTEMAS

PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN SISTEMAS

TEMA:

Estudio comparativo sobre el uso de los asistentes virtuales Siri y Cortana, basado en distintas plataformas de desarrollo.

EGRESADA:

Navarrete Masalema Ana Gabriela

TUTOR:

Ing. Enrique Ismael Delgado Cuadro, MTE

AÑO 2021

RESUMEN

Los asistentes de voz, conocidos también como interfaces de usuario naturales, son agentes de software que pueden interpretar el habla humana y responder a través de voces sintetizadas. Una interfaz de usuario natural es un sistema para la interacción persona-computadora que el usuario opera a través de acciones intuitivas "invisibles". El objetivo de estas interfaces es ocultar la complejidad del sistema incluso si el usuario tiene experiencia o las interacciones son complejas. Las interfaces de usuario naturales se están volviendo populares. Este caso de estudio presenta los resultados de un análisis de dos asistentes personales inteligentes, Siri y Cortana, en las siguientes dimensiones: la exactitud de sus respuestas, qué tan naturales se sienten las respuestas para los usuarios, la compatibilidad y la seguridad. resulta fundamental centrarnos en la compatibilidad con los dispositivos y con las plataformas y servicios de streaming. Poco sirve decantarse por el sistema de Apple y su calidad de sonido, si no usamos Apple Music o si queremos controlar la televisión y no tenemos Apple TV. Cortana está por delante al contar con un ecosistema más amplio y nutrido que el de Siri. Por otro lado, Apple ofrece un ecosistema más cerrado, sólido y seguro, debido a que todo está certificado y aprobado por la empresa.

Palabras clave

estudio, comparativo, asistente, siri, cortana

ABSTRACT

Voice assistants, also known as natural user interfaces, are software agents that can interpret human speech and respond through synthesized voices. A natural user interface is a system for human-computer interaction that the user operates through intuitive "invisible" actions. The goal of these interfaces is to hide the complexity of the system even if the user is experienced or the interactions are complex. Natural user interfaces are becoming popular. This case study presents the results of an analysis of two smart personal assistants, Siri and Cortana, on the following dimensions: the accuracy of their responses, how natural the responses feel to users, compatibility, and security. It is essential to focus on compatibility with devices and with streaming platforms and services. It is of little use to opt for the Apple system and its sound quality, if we do not use Apple Music or if we want to control the television and we do not have Apple TV. Cortana is ahead by having a broader and more nourished ecosystem than Siri. On the other hand, Apple offers a more closed, solid and secure ecosystem, because everything is certified and approved by the company.

Keywords

study, comparative, assistant, siri, cortana

INTRODUCCIÓN

Los asistentes de voz, conocidos también como interfaces de usuario naturales, son agentes de software que pueden interpretar el habla humana y responder a través de voces sintetizadas. Siri de Apple, Alexa de Amazon, Cortana de Microsoft y el Asistente de Google son los asistentes de voz más populares y están integrados en teléfonos inteligentes o altavoces domésticos dedicados. Los usuarios pueden hacer preguntas a sus asistentes, controlar los dispositivos de automatización del hogar y la reproducción de medios a través de la voz, y administrar otras tareas básicas como el correo electrónico, las listas de tareas pendientes y los calendarios con comandos verbales.

Una interfaz de usuario natural es un sistema para la interacción persona-computadora que el usuario opera a través de acciones intuitivas "invisibles". El objetivo de estas interfaces es ocultar la complejidad del sistema incluso si el usuario tiene experiencia o las interacciones son complejas. A pesar de que los asistentes personales inteligentes ahora son la corriente principal, evaluar a estos asistentes representa un desafío debido a la gran variedad y cantidad de tareas que realizan. Por ejemplo, los asistentes que se encuentran en el teléfono inteligente promedio admiten una amplia gama de tareas, como comandos de voz, búsqueda web, chat y varias otras. Debido a la cantidad de tareas que usan comandos de voz, los estudios que intentan medir la efectividad de estos asistentes o compararlos.

Una de las interfaces de usuario naturales más comunes en la actualidad son las interfaces activadas por voz. Se supone que las interfaces de usuario naturales deben ser utilizadas por humanos de una manera muy lógica. Sin embargo, la carrera para implementarlas basadas en voz por parte de la industria ha tenido un gran impacto en la naturalidad de tales interfaces. Las interfaces de usuario naturales se están volviendo populares. Este caso de estudio presenta los

resultados de un análisis de dos asistentes personales inteligentes, Siri y Cortana, en dos dimensiones: la exactitud de sus respuestas y qué tan naturales se sienten las respuestas para los usuarios.

Para el presente estudio de caso se utilizó la metodología cualitativa, porque se presenta una prueba funcional y de usabilidad de los asistentes personales activados por voz es decir Apple Siri, Microsoft Cortana. Como herramienta se utilizó una encuesta donde se explora el funcionamiento básico y las características comunes estos asistentes. También discutió algunos de los problemas de privacidad y seguridad inherentes a los asistentes de voz y los posibles usos futuros de estas tecnologías. Este estudio contó con la participación de 62 sujetos (edad promedio = 26, hombres = 36, mujeres = 26) familiarizados con la tecnología y el uso de Asistentes Personales Inteligentes. Como instrumento se utilizó un cuestionario donde todos los participantes evaluaron la calidad y la corrección de las respuestas proporcionadas por cada uno de los asistentes personales.

El presente estudio de caso se relaciona con la línea de investigación de sistemas de información y comunicación, emprendimiento e innovación conjuntamente relacionado con la sublínea de redes y tecnologías inteligentes software y hardware.

DESARROLLO

La gente ha querido hablar con las computadoras casi desde el momento en que se inventó la primera computadora. La ciencia ficción está llena de computadoras que pueden mantener una conversación. Hace solo unas décadas, la idea de mantener una conversación significativa con una computadora parecía futurista, pero la tecnología para hacer que las interfaces de voz sean útiles y ampliamente disponibles ya está aquí. (Hoy, 2018)

Varios productos de nivel de consumidor desarrollados en los últimos años han traído asistentes de voz económicos al uso diario, y se están agregando más funciones y plataformas todo el tiempo. Los usuarios pueden hacer de todo, desde hacer preguntas informativas sencillas hasta reproducir música y marcar su teléfono o encender y apagar las luces a través del control por voz. (Berdasco, 2020)

El presente estudio de caso presenta una prueba funcional y de usabilidad de algunos de los asistentes personales activados por voz es decir Apple Siri, Microsoft Cortana. Cada asistente personal respondió la misma solicitud y un grupo de ocho personas calificó cada respuesta en dos categorías: sentimiento natural y corrección. Los resultados muestran que Cortana es el asistente personal más natural. Sin embargo, también es el menos correcto. Por otro lado, Siri es el asistente personal más correcto y el menos natural.

Se explorará el funcionamiento básico y las características comunes de los asistentes de voz Siri y Cortana. También discutirá algunos de los problemas de privacidad y seguridad inherentes a los asistentes de voz y los posibles usos futuros de estas tecnologías. Dado que los asistentes de voz se utilizan más ampliamente, los bibliotecarios querrán familiarizarse con la tecnología, que tiene el potencial de ser un mecanismo en la gestión de procesos empresariales y de producción.

En pocas palabras, los asistentes de voz son la realización del sueño de ciencia ficción de interactuar con nuestras computadoras hablando con ellas. Siri de Apple, Cortana de Microsoft, Alexa de Amazon y el Asistente de Google son todos agentes de software que se ejecutan en dispositivos de altavoz o teléfonos inteligentes especialmente diseñados. El software escucha constantemente una palabra clave para despertarlo. Una vez que escucha esa palabra clave, graba la voz del usuario y la envía a un servidor especializado, que la procesa e interpreta como un comando. Dependiendo del comando, el servidor proporcionará al asistente de voz la información apropiada para ser leída al usuario, reproducir los medios solicitados por el usuario o completar tareas con varios servicios y dispositivos conectados.

La cantidad de servicios que admiten comandos de voz está creciendo rápidamente y los fabricantes de dispositivos de Internet de las cosas también están incorporando el control de voz en sus productos. El asistente Siri de Apple ha existido por más tiempo, lanzado como una aplicación independiente en 2010 e incluido en iOS en 2011. Microsoft siguió después con Cortana en 2013. Amazon lanzó Alexa con su altavoz doméstico conectado a Echo en 2014, y se anunció el Asistente de Google en 2016 junto con su altavoz Home y también está integrado en la aplicación de Google para teléfonos inteligentes basados en Android. (Jiménez & Sánchez, 2015)

Cada asistente tiene sus propias características únicas, pero las funciones principales son las mismas. Los asistentes de voz se diferencian de las tecnologías activadas por voz anteriores en que pueden responder a un número mucho mayor de comandos y preguntas. Esto se debe a que siempre están conectados a Internet; cada interacción se envía de vuelta a un sistema informático central que analiza los comandos de voz del usuario y proporciona al asistente la respuesta adecuada. Los primeros dispositivos activados por voz dependían de un conjunto más pequeño de comandos y respuestas "integrados". Los avances recientes en el procesamiento

del lenguaje natural, también conocido como lingüística computacional, han permitido a los asistentes de voz crear respuestas significativas rápidamente. (Pérez, 2020)

A medida que las computadoras personales se han vuelto más baratas y más potentes, y la gente ha creado más y más texto en línea para analizar, los científicos han utilizado ese texto para capacitar a los asistentes de voz para que escuchen y respondan a nuestras solicitudes de manera más natural y significativa. Los asistentes de voz pueden analizar las solicitudes expresadas de diferentes formas e interpretar lo que es más probable que desee el usuario.

Por ejemplo, para pedirle Siri que recuerde dónde estacionó su automóvil, un usuario puede decir cualquiera de las siguientes frases: "Recuerda dónde estacioné", "Estacioné aquí", "Dejé el automóvil en la calle 6, "O" el auto está en el estacionamiento sur "obtendrán un resultado similar. Siri recordará dónde estacionó el automóvil el usuario y, cuando se le pregunte más tarde, podrá responder en consecuencia. El usuario puede hacer preguntas de una manera igualmente natural; preguntar "dónde estacioné", "dónde dejé el auto" o "¿recuerdas dónde estacioné?", todos desencadenan la respuesta esperada. El procesamiento del lenguaje natural evita la frustración del usuario de los sistemas de reconocimiento de voz anteriores, que requerían frases y patrones específicos para funcionar correctamente.

Aunque cada asistente de voz disponible actualmente tiene características únicas, comparten algunas similitudes y pueden realizar las siguientes tareas básicas:

- Enviar y leer mensajes de texto, hacer llamadas telefónicas y enviar y leer correos electrónicos;
- Mensajes;
- Responder consultas informativas básicas ("¿Qué hora es? ¿Qué tiempo hace? ¿pronóstico? ¿Cuántas onzas hay en una taza?";

- Establecer temporizadores, alarmas y entradas de calendario;
- Establecer recordatorios, hacer listas y hacer cálculos matemáticos básicos;
- Controlar los dispositivos habilitados para Internet de las cosas, como termostatos, luces, alarmas y cerraduras; y
- Contar chistes e historias.

Además de estas tareas, los asistentes de voz pueden agregar otras funciones, a menudo llamadas "habilidades", que amplían sus capacidades al interactuar con otros programas a través de comandos de voz. Estas habilidades son creadas por desarrolladores externos, de manera similar a la forma en que se desarrollan las aplicaciones para teléfonos inteligentes. Por ejemplo, decirle al asistente "Buenos días" podría iniciar una serie de acciones diseñadas para acelerar la rutina matutina del usuario: encender la cafetera, leer las noticias y los eventos del calendario, abrir la puerta del garaje y arrancar el automóvil, y luego cerrar las puertas y ajustando el termostato después de que los ocupantes hayan abandonado la casa durante el día.

Todos los dispositivos necesarios para realizar estos de forma remota están disponibles hoy; un asistente de voz simplemente proporciona el puente que permite a los usuarios emitir comandos verbalmente en lugar de a través de una aplicación. Los asistentes de voz están disponibles en una amplia variedad de plataformas de hardware. (Martínez, 2016)

Apple acaba de entrar en el mercado de altavoces domésticos, con el anuncio de su dispositivo HomePod habilitado para Siri, programado para estar disponible en diciembre de 2017. Microsoft se ha centrado en incorporar Cortana en PC y teléfonos con Windows 10 y recientemente se asoció con Harmon Kardon para desarrollar una Cortana como un altavoz doméstico habilitado. (Assef, 2019)

Los asistentes también están disponibles en la mayoría de las plataformas de teléfonos inteligentes; Microsoft y Amazon están trabajando juntos para llevar Cortana a los dispositivos de Amazon y Alexa. para PC. Apple ha mantenido a su asistente fuera de los dispositivos que no son iOS, pero Siri está disponible en todos los dispositivos Apple, incluidos iPhones, Macbooks, iPads y Apple Watch.

A medida que el mercado de asistentes de voz se estabilice, parece probable que haya una integración adicional y que los conjuntos de funciones de los asistentes de voz principales se vuelvan similares. Apple también puede convertirse en un competidor más con el lanzamiento de HomePod a fines de 2017 y la adición de más productos para el hogar conectado de la marca Apple. Es poco probable que Microsoft gane mucha tracción, ya que su participación en el mercado de teléfonos inteligentes es insignificante y carecen de un producto atractivo para el hogar. (Soto, 2019)

Si bien los asistentes de voz tienen características interesantes y útiles, también plantean varios problemas únicos. Uno de los principales problemas con estos dispositivos activados por voz es la seguridad. Cualquiera que tenga acceso a un dispositivo activado por voz puede hacerle preguntas, recopilar información sobre las cuentas y los servicios asociados con el dispositivo y pedirle que realice tareas. Esto representa un riesgo de seguridad importante porque estos dispositivos leerán el contenido del calendario, los correos electrónicos y otra información altamente personal. En un caso reportado, un hombre descubrió que el iPad en su sala de estar abriría la puerta de entrada para cualquiera que se quedara afuera y le pidiera a Siri que los dejara entrar. (Serrano, 2016)

Apple también le está enseñando a Siri a reconocer la voz de un usuario. De forma predeterminada, cualquier persona con acceso de voz al dispositivo puede solicitar artículos

utilizando la cuenta de las tiendas en línea. Hay opciones para establecer una contraseña de voz para confirmar las compras, y todos los productos se enviarán a la dirección del propietario registrada, pero aún existe la posibilidad de que usuarios malintencionados compren productos en la cuenta del propietario. Los miembros del hogar también podrían hacer compras no autorizadas, como la niña de seis años que se ordenó una casa de muñecas y cuatro libras de galletas de azúcar a través de Siri. (Jáuregui, 2021)

Los asistentes de voz también son vulnerables a varios otros ataques. Los investigadores han demostrado recientemente que los asistentes responderán a comandos inaudibles emitidos en frecuencias ultrasónicas. Esto permitiría a un atacante acercarse a una víctima, reproducir el comando ultrasónico y el dispositivo de la víctima respondería. También existe la posibilidad de que este tipo de ataque se incruste en los medios de difusión. (De los Santos, 2017)

Un anuncio de producto en la televisión que contenga comandos de ultrasonido integrados para agregar el artículo a su carrito de compras podría estimularlo a comprarlo. La idea parece descabellada, pero una serie de noticias sobre la niña mencionada anteriormente que se ordenó una casa de muñecas causó problemas a muchas personas. El presentador de noticias pronunció las palabras: "Siri, pídemme una casa de muñecas" mientras informaba la historia, y cientos de Siri en las casas de los espectadores respondieron.

La privacidad es otra preocupación importante para los usuarios de asistentes de voz. Por su propia naturaleza, estos dispositivos deben estar escuchando en todo momento para que puedan responder a los usuarios Apple. Microsoft insiste en que sus dispositivos no están grabando a menos que los usuarios digan el comando para despertar al asistente, pero ha habido al menos un caso en el que un dispositivo que funcionaba mal estaba grabando en todo momento y enviando esas grabaciones a los servidores.

Los asistentes de voz tienen el potencial de cambiar radicalmente la forma en que los usuarios interactúan con las computadoras. Para muchos usuarios, la capacidad de leer y escribir es una barrera para acceder a la información. Los asistentes de voz pueden cerrar la brecha de información para esos usuarios. Investigaciones recientes han demostrado que los asistentes de voz pueden beneficiar a quienes padecen demencia al proporcionar una voz siempre presente que puede responder las mismas preguntas una y otra vez sin perder la paciencia y ofrecer aliento cuando sea necesario. (Brasero, 2020)

Para otros, leer las instrucciones que proporciona su médico puede resultar difícil. Incorporar estas capacidades en las tecnologías de consumo actualmente disponibles sería mucho más rentable que un dispositivo especialmente diseñado, y muchos usuarios ya se sentirían cómodos operando estos dispositivos. Los asistentes de voz también podían leer libros y otros documentos de formato largo a los usuarios. Aunque todavía suenan algo robóticos, las cualidades vocales de los asistentes de voz están mejorando rápidamente. Una vez que mejoren lo suficiente como para no ser desagradables, cada libro podría convertirse en un audiolibro.

Los asistentes de voz también tienen el potencial de revolucionar la traducción. Los usuarios inician la aplicación pidiendo al asistente que les ayude a hablar un idioma. El teléfono del usuario captura el audio hablado por la otra parte, lo transmite a los servidores de traducción y reproduce la versión traducida en el auricular del usuario. Cuando el usuario quiere que se traduzca su discurso, presiona un botón en el auricular, que envía el audio al servidor, donde se traduce y se transmite al oyente a través del altavoz del teléfono del usuario. Si bien las traducciones a través de asistentes son notoriamente propensas a errores, en particular para el lenguaje médico o coloquial, los resultados casi instantáneos deberían hacerla útil para conversaciones simples. (Saorín & Pastor, 2020)

Las siglas de IFTTT significan "IF This, Then That", lo que traducido al español significaría "Si esto, entonces aquello". Con este nombre, se refiere a que es una página con la que puedes crear y programar acciones entre diferentes aplicaciones, de manera que, si haces o pasa determinada cosa con una aplicación, entonces automáticamente realiza una acción con ella. Se trata de unos de los portales más populares a la hora de sacar un mayor partido con tus redes sociales y los diferentes servicios online utilizados, ya que permite crear interacciones y tareas automáticas con las que vas a poder ahorrar bastante tiempo. Permite crear automatismos, de manera que cuando realizas una acción con alguna de tus aplicaciones, automáticamente se realizan las otras tareas que hayas asignado. (Ordaz, 2020)

Los asistentes de voz también pueden ser útiles para la promoción y gestión de bibliotecas. Ya existen herramientas disponibles que permiten a las bibliotecas crear habilidades para asistentes de voz que enumeran eventos en la biblioteca en los calendarios de la comunidad local. Codificar funciones adicionales que permitirían a los usuarios escuchar horarios y anuncios, renovar sus artículos, escuchar qué nuevos artículos están disponibles o programar una consulta con un bibliotecario sería relativamente simple.

Los asistentes de voz también podrían programarse fácilmente para actuar como guías turísticos virtuales en galerías más pequeñas o espacios de exhibición. Los clientes pueden pedirle al asistente que les cuente sobre una exhibición, y el asistente puede leer los comentarios preparados. Las bibliotecas con un enfoque tecnológico pueden considerar prestar estos dispositivos y brindar capacitación básica para que los usuarios también puedan experimentar con estos dispositivos en sus hogares.

Siri y Cortana son algunos de los asistentes personales inteligentes más comunes y más utilizados disponibles en el mercado. Desde la perspectiva académica, son interfaces de usuario

naturales basadas en el habla. En el dominio de Interacción persona computadora, los asistentes basados en el habla son sistemas que los usuarios operan a través de acciones intuitivas relacionadas con el comportamiento humano natural a través de instrucciones de voz.

Hoy en día, la mayoría de los fabricantes de dispositivos inteligentes tienen su propia aplicación de asistente de voz. Sin embargo, los diseños varían significativamente y sus funcionalidades son muy diferentes. La carrera de los fabricantes por lanzar nuevas versiones de sus asistentes personales activados por voz ha tenido un gran impacto en sus productos, especialmente en su simplicidad y usabilidad. Un ejemplo de esta afirmación es que cuando los asistentes personales activados por voz se liberaron por primera vez, su comportamiento era limitado y los comandos eran muy rígidos y estructurados. Posteriormente, se realizaron varios esfuerzos para crear interfaces más naturales. Sin embargo, algunos de los dispositivos actuales volvieron a los comandos rígidos y estructurados.

A continuación, se describe una evaluación de dos de los asistentes personales inteligentes más populares.

Cortana

Cortana es un asistente personal inteligente creado por Microsoft para dispositivos con Windows (tanto para PC como para dispositivos móviles). Cortana puede establecer recordatorios, reconocer la voz natural sin la necesidad de ingresar con el teclado y responder. Algunos de los comandos promocionales de Cortana incluyen: "¿Cómo está el clima?", "Llamar a Sarah", "¿Qué son cinco millas en kilómetros?", "¿Cuál es la definición de 'mesa'?"

¿Cómo fue desarrollado Cortana?

El desarrollo de Cortana en realidad comenzó en 2009, bajo la dirección de Zig Serafin y Larry Heck, quienes eran ingenieros que trabajaban con Microsoft Research y el equipo de Microsoft Speech. Heck y Serafin se propusieron crear un asistente digital que no solo brindaría asistencia con las tareas del día a día, sino que también aprendería sus hábitos y respondería de manera anticipada.

Para Heck, el primer lanzamiento ofreció una muestra de lo que tenía en mente. Con el tiempo, Heck quería que Cortana interactuara de una manera cada vez más anticipada y natural. Si bien no es una verdadera IA, Cortana exhibe rasgos similares a los humanos cada vez más reconocibles, y eso no es casualidad. Esto es lo que Heck tenía que decir en 2014.

Las tecnologías básicas para un asistente personal virtual incluyen reconocimiento de voz, procesamiento de lenguaje semántico / natural, modelado de diálogo entre humanos y máquinas, y generación de lenguaje hablado. Cada área contiene una serie de problemas de investigación que Microsoft Research ha abordado a lo largo de los años. De hecho, hemos sido pioneros en los esfuerzos en cada una de esas áreas.

Cortana aprovecha el motor de búsqueda web de Bing para proporcionar información de Internet, mientras utiliza algoritmos internos de aprendizaje automático desarrollados por Microsoft Research. Cada vez más empresas están invirtiendo en tecnología de inteligencia artificial, incluidas Google y Facebook, ya que todas las principales empresas de tecnología compiten por una mano en el próximo gran cambio de paradigma tecnológico.

Microsoft ha construido Cortana intencionalmente para escalar a todos los diferentes dominios. Tener una visión a largo plazo significa que tenemos una arquitectura a largo plazo. El objetivo es respaldar todo tipo de interacción humana, ya sea de habla, texto o gestos, en los

dominios de la información y la función, y hacer que sea tan fácil como una conversación natural.

En 2016, Microsoft anunció que los desarrolladores podrían conectarse con Cortana a través de una nueva plataforma de bots, además de conectarse directamente para interacciones naturales. Ahora puede hacer llamadas de Skype, buscar y reproducir aplicaciones de música de terceros e incluso llamar a taxis de Uber, utilizando los comandos de lenguaje natural de Cortana.

Cortana se programó para su primera salida con un combo de altavoz y micrófono dedicado en 2017 con el altavoz Harman Kardon "Evoke". Y dio el salto a Android, iOS y Xbox One a lo largo de 2015 y 2016. Con Cortana en cientos de millones de dispositivos con Windows 10 (al momento de la redacción), ha contribuido a un gran crecimiento en el motor de búsqueda Bing, que históricamente ha tenido problemas. para competir contra el dominio de Google.

Microsoft ofrece continuamente más capacidades a Cortana a través del Cortana Devices SDK y el programa Cortana Skills , que cuenta con el apoyo de docenas de desarrolladores y proveedores externos, incluidos Spotify, Fitbit, Expedia, Dark Sky, OpenTable y otros. Puede esperar que Cortana continúe creciendo en todos los dispositivos y plataformas y entregue nuevas habilidades y capacidades a través de nuestros socios externos, tanto en el espacio empresarial como en el espacio del consumidor, dijo Clayton. Esperamos poder ofrecer más productividad y escenarios de trabajo a Cortana en un futuro próximo.

Siri

Siri es un asistente personal para dispositivos Apple. Lanzado originalmente en 2010, es el asistente personal inteligente más antiguo de los más utilizados. Los usos más comunes de

Siri son la navegación web y el dictado. Actualmente, Siri se encuentra actualmente en un proceso de rediseño para agregar nuevas funcionalidades y adaptarlo a nuevos dispositivos. Algunos de los comandos promocionales para Siri incluyen: "Muestre mis fotos en agosto pasado", "¿Qué películas se están proyectando hoy?", "Busque videos que tomé en la fiesta de cumpleaños de María", "Envíe un mensaje de texto a Pete con un signo de exclamación sonriente y nos vemos pronto", " Lee mi último correo electrónico ", " Encuentra una mesa para cuatro esta noche en Babahoyo "

¿Cómo se desarrolló Siri?

Apple no desarrolló completamente Siri por sí mismos. Comenzó a partir de una iniciativa de inteligencia artificial en 2003 financiada por DARPA (Agencia de Proyectos de Investigación Avanzada de Defensa) y dirigida por una filial de la Universidad de Stanford, SRI International. Su objetivo era crear un programa que ayudara al personal militar con el trabajo de oficina y la toma de decisiones, lo que resultó en CALO (Asistente cognitivo que aprende y organiza) que aprende de sus usuarios y de la gran cantidad de datos disponibles. Se utilizó para organizar y programar reuniones, así como para proporcionar los documentos necesarios para los participantes. CALO incluso tomó decisiones en situaciones en las que alguien importante no pudo asistir a la reunión, puede cancelar la reunión y reprogramar o continuar la reunión según lo planeado si su presencia no es necesaria. Vanguard fue otro proyecto desarrollado por SRI International que era un prototipo que funcionaba muy bien para teléfonos inteligentes, pero que no tenía las capacidades de CALO.

Se formó una empresa de nueva creación que tenía empleados exalumnos de la NASA y Google para combinar Vanguard y CALO. Esta empresa se llamó Siri y la aplicación que crearon se llamó Siri Assistant. Esta versión tomaría preguntas de los usuarios a través de la

voz o pulsaciones de teclas, enviaría esos datos a un servidor remoto para traducirlos y luego buscaría en varios sitios web. Siri Assistant tenía un tipo de humor único y una predilección por las respuestas bocas inútiles. Cuando Apple adquirió la compañía Siri, eliminó algunas de sus características, como su humor y lenguaje inapropiado, y el acceso a sitios web de la competencia para priorizar sus propios servicios, pero ganó capacidades multilingües, características específicas de iPhone y le dio la voz que queríamos. están familiarizados con la actualidad.

Con las variaciones y matices sutiles que tenemos en nuestro discurso hay una infinidad de formas de construir frases. Imitar la forma compleja en que los humanos comprenden el habla en un software de programación es una empresa monumental. Voz de teléfono inteligente. Como resultado, los desarrolladores de Siri, junto con la compañía de software de dictado Nuance Communications, han programado su software de reconocimiento de voz para interpretar comandos y preguntas a través de una serie de pasos que permiten a las personas interactuar con Siri de la manera más humana posible.

Al recibir su solicitud, Siri registra las frecuencias y ondas sonoras de su voz y las traduce a un código. Siri luego desglosa el código para identificar patrones, frases y palabras clave particulares. Estos datos se ingresan en un algoritmo que analiza miles de combinaciones de oraciones para determinar qué significa la frase ingresada. Este algoritmo es lo suficientemente complejo como para trabajar con modismos, homófonos y otras expresiones literarias para determinar el contexto de una oración. Una vez que Siri determina su solicitud, comienza a evaluar qué tareas deben realizarse, determinando si se puede acceder a la información necesaria desde los bancos de datos del teléfono o desde los servidores en línea. Luego, Siri puede elaborar oraciones completas y coherentes relevantes para el tipo de pregunta o comando solicitado.

Este estudio contó con la participación de 62 sujetos (edad promedio = 26, hombres = 36, mujeres = 20) familiarizados con la tecnología y el uso de Asistentes Personales Inteligentes. Durante la ejecución de la evaluación, todos los participantes se reúnen en el mismo lugar. Todos los participantes evaluaron la calidad y la corrección de las respuestas proporcionadas por cada uno de los asistentes personales inteligentes respondiendo las siguientes dos preguntas: "¿Qué tan buenas fueron las respuestas?" y "¿Cuán correctas fueron las respuestas?".

La Tabla 1 describe los dispositivos utilizados para cada asistente personal y la Figura 1 muestra los dispositivos.

Asistente personal	Dispositivo
Siri	iPhone 7 S
Cortana	Hp Pavilion Dev 7

Tabla 1. Dispositivos utilizados. Elaboración propia

Hay muchas clasificaciones de funcionalidades de asistente. Para definir el procedimiento de evaluación usaremos la siguiente clasificación: asistente de compras y compras, asistente de cuidado, asistente de viajes y entretenimiento y asistente administrativo. Todos los participantes escucharon las respuestas y dieron una respuesta usando una escala Likert de 5 puntos para la naturalidad de la respuesta y la corrección.

Las llamadas "escalas Likert" son instrumentos psicométricos donde el encuestado debe indicar su acuerdo o desacuerdo sobre una afirmación, ítem o reactivo, lo que se realiza a través de una escala ordenada y unidimensional. Estos instrumentos suelen ser reconocidos entre los más utilizados para la medición en Ciencias Sociales. (Matas, 2018)

En la categoría de asistente de compras, se consideraron cuatro características: administrar listas de compras, ordenar productos a través de Internet, buscar restaurantes o tiendas y buscar horarios de tiendas. Para la categoría de viajes y entretenimiento, se incluyeron cinco solicitudes: actualización de deportes, información de cine, transporte, distancia y hora de llegada, proporcionando dos ubicaciones e identificando una canción.

En la categoría de asistente administrativo, se consideraron seis características: administrar múltiples temporizadores y alarmas, listas de tareas, recordatorios para la misma fecha y una fecha específica, programación de reuniones y redacción de correos electrónicos. Finalmente, para la categoría miscelánea se evaluaron 21 características: noticieros, actualizaciones de tráfico, pronóstico del tiempo, administración de dispositivos, juegos, funciones aleatorias (lanzar una moneda, números aleatorios, lanzar un dado), conversiones, medidas, moneda, calculadora de propinas, resolución de matemáticas problemas, y otras funcionalidades como (contar chistes, hacer beatboxing, responder preguntas aleatorias o contar hechos).

El beatboxing es una forma de percusión vocal que involucra principalmente el arte de imitar cajas de ritmos, usando la boca, los labios, la lengua y la voz. También puede implicar la imitación vocal y otros instrumentos musicales. (Hernández, 2018)

En la categoría de compras y compras, la mejor asistente personal fue Cortana. Sin embargo, ninguna fue capaz de comprar online. Cortana dio las respuestas más naturales en esta categoría. El asistente personal menos correcto en esta categoría fue Siri. La solicitud con mejor respuesta en todos los dispositivos fue la de encontrar un restaurante. Pero en ninguno está disponible hacer pedidos. Sin embargo, la solicitud con la respuesta más natural fueron los horarios de las tiendas. En la categoría de viajes y entretenimiento, Siri y Cortana comparten

el primer lugar en corrección y Siri fue la que tuvo las respuestas más naturales. Cortana dio las respuestas menos naturales. Además, en esta categoría, la solicitud con mejor (correcta) y más natural respuesta en general fue la actualización deportiva. En la categoría de viajes y entretenimiento, el Siri dio la mejor respuesta a la consulta de ubicación y hora de llegada. Cortana dio la peor respuesta cuando se le solicitó obtener información sobre una película.

La categoría de asistente administrativo tiene un empate en el asistente personal más correcto entre Siri y Cortana. El asistente personal más natural fue nuevamente Siri. La consulta mejor respondida fueron los temporizadores y las alarmas tanto en exactitud como en naturalidad. La mejor respuesta en general fue el manejo de temporizadores. Sin embargo, Cortana dio excelentes resultados enviando correos electrónicos y configurando recordatorios. Finalmente, Siri ganó abrumadoramente la categoría de miscelánea tanto en corrección como en naturalidad. En cuanto a la naturalidad, ganó esta función en más de un 400%. Cortana dio los peores resultados en corrección para esta categoría.

En general, al realizar el análisis estadístico no se encontraron diferencias significativas entre los sistemas. En otras palabras, la evaluación no muestra preferencia por ningún sistema, ni en naturalidad ni en corrección. Este resultado se extrajo de un análisis de todos los resultados juntos. Siri fue el dispositivo más correcto; sin embargo, Cortana fue el que tuvo las respuestas más naturales. Cortana sufrió porque no es compatible con algunas de las funciones evaluadas en esta investigación. Siri tuvo los peores resultados en naturalidad.

En el caso de Siri, al evaluar si las respuestas fueron correctas, su desempeño fue pobre ya que el 43% de los participantes considera que las respuestas están por debajo de la media (respuestas incorrectas). Este es el peor desempeño entre los cuatro asistentes. En el caso de Cortana, el 71% considera que las respuestas estuvieron por encima del promedio y el 18% las

considera promedio. En el caso de Siri y Cortana los números son considerablemente bajos, considerando que estos asistentes se utilizan para ayudar a las personas en sus actividades diarias o para resolver problemas cotidianos, y lo más importante es asegurarse de que brinden una buena comunicación y respuestas correctas.

Características técnicas de Siri

Los accesos directos son una nueva capacidad en iOS 12, permitiendo a los usuarios de iPhone y Apple Watch usar Siri para progresar a través de rutinas. Los accesos directos reemplazan la aplicación Workflow que Apple adquirió en 2017 y está diseñada para crear comandos personalizados en Siri que inician aplicaciones o combinan varias acciones de manera similar a IFTTT (If This Then That).

Apple ofrece una amplia gama de comandos de voz para interactuar con Siri, que incluyen, entre otros:

- Acciones telefónicas y de texto, como «Llamar a Sara», «Leer mis nuevos mensajes», «Configurar el temporizador para diez minutos» y «Enviar correo electrónico a mamá»
- Verificar la información básica, que incluye «¿Qué tiempo hace hoy?» Y «¿A cuántos euros está el dólar?»
- Programar eventos y recordatorios, incluidos «Programar una reunión» y «Recuérdame»
- Controla la configuración del dispositivo, como «Tomar una foto», «Activar Wi-Fi» y «Aumentar el brillo»
- Busque en Internet, incluyendo «Definir ...», «Buscar imágenes de ...» y «Buscar en Twitter ...»

- Navegación, que incluye «Llévame a casa», «¿Cómo está el tráfico de camino a casa?» Y «Encuentra indicaciones para llegar a ...»
- Entretenimiento, como «¿Qué partidos de baloncesto hay hoy?», «¿Qué películas puedo ir a ver por aquí cerca?» Y «¿Cuál es la sinopsis de ...?»
- Interactúa con aplicaciones integradas en iOS, como «Pause Apple Music» y «Like this song»

Características técnicas de Cortana

Cortana puede establecer recordatorios, reconocer la voz natural sin el requisito de entrada del teclado y responder preguntas utilizando la información del motor de búsqueda Bing (por ejemplo, condiciones climáticas y de tráfico actuales, puntajes deportivos, biografías). Cortana se integra con servicios como Foursquare para brindar recomendaciones de restaurantes y eventos locales. El 16 de febrero de 2018, Microsoft anunció que se agregarían habilidades domésticas conectadas para ecobee, Honeywell Lyric, Honeywell Total Connect Comfort, LIFX, TP-Link Kasa y Geeni. También se agregó soporte para IFTTT.

Cortana almacena información personal como intereses, datos de ubicación, recordatorios y contactos en el «Notebook». Aprovecha estos datos para conocer los patrones y comportamientos específicos de cada persona. Los usuarios pueden ver y especificar qué información se recopila para permitir cierto control sobre su privacidad. Se dice que esta función cuenta con «un nivel de control que va más allá de los asistentes comparables». Los usuarios pueden eliminar información del «Notebook».

Cortana tiene un sistema incorporado de recordatorios que, por ejemplo, se pueden asociar a un contacto específico; esto sirve para recordar al usuario temas pendientes cuando se comunique con ese contacto, posiblemente en un momento específico o cuando el teléfono

esté en una ubicación determinada. Originalmente, estos recordatorios eran únicos del dispositivo en el que estaba instalada Cortana, pero ahora Windows 10 sincroniza los recordatorios en todos los dispositivos.

En Windows Mobile y Android, Cortana es capaz de capturar notificaciones de dispositivos y enviarlas a un dispositivo con Windows 10. Esto permite que el usuario de la computadora vea notificaciones desde su teléfono en el Centro de acción de Windows 10.

	 Siri	 Cortana
TIPO DE ACCESORIOS PERMITIDOS		
CANTIDAD DE ACCESORIOS COMPATIBLES		
INTELIGENCIA DEL ASISTENTE DE VOZ		
CALIDAD SONORA		
COMPATIBILIDAD CON PLATAFORMAS DE STREAMING		
PRIVACIDAD Y SEGURIDAD		

Figura 1. Comparativa entre Siri y Cortana. (Brasero, 2020)

Compatibilidad

Siri sólo está disponible en dispositivos creados por Apple, mientras que Cortana está optando por integrarse a otras aplicaciones, incluidos teléfonos inteligentes, tabletas, ordenadores y dispositivos portátiles. Parece poco probable que Microsoft continúe

desarrollando Windows 10 Mobile. Ahora se centrará en proporcionar el servicio de Cortana con iOS y Android.

Asistente de voz	Dispositivos disponibles
Siri	IOS 9 o superior MacOS posterior a 10.14 Mojave
Cortana	Windows 10 iOS 9.0 o superior Android 4.4 o superior

Tabla 2. Dispositivos compatibles para Siri y Cortana

Seguridad

La privacidad es una de las categorías en las que Apple destaca positivamente, ya que protege al usuario de forma más segura y respetuosa. La diferencia, con respecto a Cortana, es que los datos de las grabaciones registradas no se asocian a un perfil personal. De este modo, además de proteger la privacidad del usuario, también significa que Apple no utiliza esa información para dirigirnos publicidad, vendernos productos o ir mejorando el sistema de hogar inteligente.

Por otro lado, Siri controla los dispositivos llevando a cabo una solicitud a su sistema a través de un identificador aleatorio, que no puede vincularse a un usuario específico, mientras que sus homólogos almacenan en la nube las grabaciones de voz durante 18 meses. Tras este periodo, las grabaciones y sus correspondientes transcripciones se eliminan de forma automática. No obstante, se puede eliminar en cualquier momento esta información de forma manual. Apple también supera, en materia de seguridad, con respecto a Cortana, ya que cualquier dispositivo que quiera conectarse a HomeKit requiere de un cifrado especial.

Podría decirse que la diferencia más marcada entre estos dos asistentes virtuales es con qué hardware y plataformas pueden integrarse. Siri puede funcionar bien con bocinas Homepod, auriculares AirPods y en dispositivos como el iPhone y el iPad, pero no puedes usarlo en tu dispositivo Android o PC con Windows. Cortana por su parte, cuenta con una amplia gama de soporte de plataforma de software y hardware. Se integra en todos los dispositivos Windows 10, incluidos todas las PC con este sistema operativo, el Windows 10 Mobile (ahora desaparecido) y la consola Xbox One. Cortana también vive en un par de dispositivos inteligentes, incluida la bocina Harmon Kardon Invoke, y existen aplicaciones de Cortana cada vez más poderosas e importantes para iOS y Android.

CONCLUSIONES

La complejidad y precisión de la tecnología de reconocimiento de voz y el software de asistente de voz han crecido exponencialmente en los últimos años. Los productos de asistente de voz actualmente disponibles de Apple y Microsoft permiten a los usuarios hacer preguntas y emitir comandos a las computadoras en lenguaje natural. Hay muchos usos posibles de esta tecnología en el futuro, desde la automatización del hogar hasta la traducción, el acompañamiento y el apoyo a las personas mayores. Sin embargo, también existen varios problemas con los productos de asistente de voz disponibles actualmente.

La ciencia ficción ha mostrado el concepto de asistentes inteligentes a través de películas, programas de televisión y libros. Las nuevas tecnologías y algoritmos han permitido el desarrollo de asistentes personales para uso comercial desde hace algunos años. Los asistentes disponibles proporcionan una amplia variedad de funcionalidades. Sin embargo, esta variedad depende de la implementación del asistente personal y su propósito.

Existe una oportunidad de mejora en todos los dispositivos probados. No se observó ninguna tendencia. Todos los dispositivos mostraron al menos una funcionalidad con respuestas incorrectas o poco naturales. Sin embargo, el Asistente de Google fue el mejor en la categoría de varios y el peor en las otras categorías. Una característica notable de Cortana fue la naturalidad de responder algunas preguntas. El tono y el ritmo de la voz femenina que utiliza Siri expresa sorpresa, suspenso y alegría. Cortana no siempre ofrecen estas funciones. Sin embargo, algunas respuestas fueron catalogadas como naturales usando estos dispositivos.

Los controles de privacidad y seguridad deberán mejorarse antes de que los asistentes de voz se utilicen para cualquier cosa que requiera confidencialidad. Los bibliotecarios deben monitorear estos productos y estar listos para brindar asistencia a sus usuarios con estos

dispositivos. También deben explorar las posibilidades de proporcionar materiales de biblioteca a través de asistentes de voz a medida que madura la tecnología. El potencial de los asistentes personales inteligentes podría probarse en áreas inexploradas como asesoramiento, marketing, aprendizaje y ventas. Además, la combinación de estos dispositivos y tecnologías con robots, centros de datos y técnicas de aprendizaje automático brinda nuevas oportunidades.

En este sentido, resulta fundamental centrarnos en la compatibilidad con los dispositivos y con las plataformas y servicios de streaming. Por tanto, de poco sirve decantarse por el sistema de Apple y su calidad de sonido, si no usamos Apple Music o si queremos controlar la televisión y no tenemos Apple TV. Así que, Cortana está por delante al contar con un ecosistema más amplio y nutrido que el de Siri. Y está presente en distintas plataformas y ciertos tipos de dispositivos, es sólo cuestión de tiempo que vayan lanzando nuevas incorporaciones para equipararse completamente a iOS y Android. Por otro lado, Apple ofrece un ecosistema más cerrado, sólido y seguro, debido a que todo está certificado y aprobado por la empresa, de acuerdo con la experiencia que busca proporcionar a sus usuarios.

En cuanto a economía, para usar cualquiera de los dos asistentes, se requiere adquirir un dispositivo, en el caso de Cortana, se puede optar por un ordenador con Windows 10 el cual el más económico es de unos \$300. Por otro lado, para poder utilizar Siri se debe adquirir un iPhone; del cual el más económico oscila entre los \$600, tomando en cuenta que este último aporta al usuario a más de seguridad, le otorga portabilidad.

BIBLIOGRAFÍA

- Assef, J. (2019). *HUMANOS Y (ro)BOTS ¿Sistemas interaccionales?* (Tesis de maestría) Santiago de Chile.
- Berdasco, A. (2020). *Evaluación de asistentes inteligentes por voz con base en la calidad de las respuestas.* . (Tesis de maestría) Universidad de Costa Rica, Ciudad Universitaria Rodrigo Facio.
- Brasero, A. M. (2020). *Desarrollo de skills accesibles para asistentes de voz comerciales.* (Tesis de maestría) Universidad Autónoma de Madrid, Madrid.
- De los Santos, S. (2017). *Experiencia de uso de asistentes de voz sin GUI en personas mayores.* (Tesis de maestría) Universidad Internacional de Rioja, Madrid.
- Hernández, W. (2018). Aproximación dialógica a la música y la radio comunitaria. *Nómadas*.
- Hoy, M. (2018). Alexa, Siri, Cortana, and More: An Introduction to. *Medical Reference Services Quarterly*, 81-88.
- Jáuregui, I. (2021). Oye Siri. *JONNPR*, 1240-44.
- Jiménez, P., & Sánchez, J. (2015). *DE ELIZA A SIRI: LA EVOLUCIÓN* . (Tesis de pregrado) Escuela Politécnica Superior Villanueva de la Cañada, Madrid.
- Martínez, E. (2016). Prototipo de Asistente Virtual para la Carta de Servicios del Servicio de Informática de la Universidad de Jaén. *Procesamiento del Lenguaje Natural*, 319-320.
- Matas, A. (2018). Diseño del formato de escalas tipo Likert: un estado de la cuestión. *Revista electrónica de investigación educativa*, 38-47.

- Ordaz, T. (2020). Valoración de estrategias de construcción del conocimiento en los entornos personales de aprendizaje. *Apertura*, 6-21.
- Pérez, M. (2020). Análisis del uso de asistentes virtuales en el aula como recurso complementario en la práctica docente. *Enseñanza y Aprendizaje de Ingeniería de Computadore*, 61-66.
- Sánchez, J. (2015). *Clima Organizacional*. Madrid: Editorial Academica Espanola.
- Saorín, T., & Pastor, J. (2020). Del mobile-first al data-first: schema. org, búsquedas zero-click y la incertidumbre sobre los asistentes de voz. *Anuario ThinkEPI*, 1-10.
- Serrano, J. (2016). Tendencias tecnológicas en internet: hacia un cambio de paradigma. *El profesional de la información*, 843-850.
- Soto, H. (2019). La inteligencia artificial: al alcance de la mano. *Convicciones*, 69-73.

ANEXO I.

Resultados de la evaluación de las funciones de los asistentes personales por función.

Abreviaturas: S = Siri, C = Cortana.

Característica	La mejor		Peor	
	Corrección	Naturalidad	Corrección	Naturalidad
Asistente de compras				
Introducciones	S		C	
Gestionar en línea	C			
Comprar en línea			S, C	S, C
Encontrar restaurantes	C	C		S
Horarios de la tienda	C	C	S	
Asistente de viajes y entretenimiento				
Actualizaciones deportivas	S	S, C		
Información de la película	C	S		C
Información de transporte	S			C
Tiempo de transporte				S
Identificar una canción	C	C		
Asistente administrativo				
Alarmas y temporizadores	S	S	C	C
Administrar listas de tareas				
Recordatorios	S, C	S		
Recordatorios en una cita	C	S		
Programar una reunión	C	C		
Redactar un correo electrónico	S	S		
Misceláneos				
Noticias locales	C		S	S
Actualizaciones de tráfico	C	C		
Pronóstico del tiempo	S, C	S		C
Gestión de dispositivos				
Juegos			S	S
Aleatorio (moneda)			C	C
Aleatorio (dados)	S, C			S
Preguntas al azar			S	S
Hechos aleatorios	S		C	
Conversiones		C	S, C	S
Mediciones	C	S		
Divisa	S		C	C
Información de vacaciones			C	C
Calcular propinas				
Problemas de matemáticas	S, C			
Zonas horarias			C	C
Chistes		C	S	S
Beatbox				
Traducción				
Definiciones de palabras			S, C	S, C

Pregunta	Asistente	Cortana	Siri
Asistente de compras			
¿Qué tan buenas fueron las respuestas?			
¿Qué tan correctas fueron las respuestas?			
Asistente de viajes y entretenimiento			
¿Qué tan buenas fueron las respuestas?			
¿Qué tan correctas fueron las respuestas?			
Asistente administrativo			
¿Qué tan buenas fueron las respuestas?			
¿Qué tan correctas fueron las respuestas?			
Misceláneos			
¿Qué tan buenas fueron las respuestas?			
¿Qué tan correctas fueron las respuestas?			

Respuestas

- (1) muy deficiente,
- (2) deficiente
- (3) bueno
- (4) muy bueno
- (5) excelente.